



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 144963

TO: Bao-Qun Li
Location: rem/3d24/3c18
Art Unit: 1648
Tuesday, February 22, 2005

Case Serial Number: 09/664363

From: Edward Hart
Location: Biotech-Chem Library
REM-1A55
Phone: 571-272-2512

edward.hart@uspto.gov

Search Notes

Examiner Li,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

This Page Blank (uspto)

STIC-Biotech/ChemLib

144963

mej

From: Li, Bao-Qun
Sent: Friday, February 11, 2005 8:10 AM
To: STIC-Biotech/ChemLib

PLEASE DO THE SEQUENCE HOMOLOGY AND INTERFERENCE OF SEQ ID NO; 4 IN APPLICATION SN.
09,664,363. THANKS.

Bao Qun Li M.D
TC 1600
Art Unit 1648
Tel. 517-272-0904
REM, 3C18
Rm. 3D24

RECEIVED
FEB 11 2005
(STIC)

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2- _____
Date Searcher Picked up: 2/11/05
Date Completed: 2/11/05
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search 1

NA Sequence: # _____
AA Sequence: # _____
Structure: # _____
Bibliographic: _____
Litigation: _____
Patent Family: _____
Other: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: STIC
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)



STIC SEARCH RESULTS

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact:*

Mary Hale, Information Branch Supervisor
Remsen Bldg. 01 D86
571-272-2507

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC-Biotech-Chem Library Remsen Bldg.



This Page Blank (uspto)


```
QY 61 GAGAGGTCAATTCCAGGTCGGGCTCAACCAATACCTGATGGGTGCGAGCTCCATGC 120
DB 61 GAGAGGTCAATTCCAGGTCGGGCTCAACCAATACCTGATGGGTGCGAGCTCCATGC 120
QY 121 GAGCCGGAACCGGATGTAGACAGTGTCTCACTTCCATGCTCAACGACCCCTCCACATGACA 180
DB 121 GAGCCGGAACCGGATGTAGACAGTGTCTCACTTCCATGCTCAACGACCCCTCCACATGACA 180
QY 181 GCAGAGACGGGCTAAGGCGACGGGCTGCGAGGGGGTCTCCCGCTCTTGGCCAGCTCTTCA 240
DB 181 GCAGAGACGGGCTAAGGCGACGGGCTGCGAGGGGGTCTCCCGCTCTTGGCCAGCTCTTCA 240
QY 241 GCTAGCAGATTGTCTGCGCTTCTCTTGAAGGCGACATATTAATGATTTCCCA 300
DB 241 GCTAGCAGATTGTCTGCGCTTCTCTTGAAGGCGACATATTAATGATTTCCCA 300
QY 301 GAGCTGACCTCATCGAGGCTCAACCTCTGTGTGGCGGATGAGATGGGGCGGACATTACC 360
DB 301 GAGCTGACCTCATCGAGGCTCAACCTCTGTGTGGCGGATGAGATGGGGCGGACATTACC 360
QY 361 CGGCTGAGTCAAGAAACAAGGATGATCTTGACCTTTTGACCCGCTCCGAGCGAG 420
DB 361 CGGCTGAGTCAAGAAACAAGGATGATCTTGACCTTTTGACCCGCTCCGAGCGAG 420
QY 421 GAGGATGACGGGGAAGTGTCCGCTCCGCGGAGATCTGCGAATTCAGAAATTTCCCA 480
DB 421 GAGGATGACGGGGAAGTGTCCGCTCCGCGGAGATCTGCGAATTCAGAAATTTCCCA 480
QY 481 CCAAGCATGCCCCGATGCGGACGCGCGGATTAACACCTTCGCTGTGAGTCTTGGAAG 540
DB 481 CCAAGCATGCCCCGATGCGGACGCGCGGATTAACACCTTCGCTGTGAGTCTTGGAAG 540
QY 541 GCCCGGACCTACGTCCTCCAGTGTGACATGAGGTGCGCACTGCACTTAAGACCCCT 600
DB 541 GCCCGGACCTACGTCCTCCAGTGTGACATGAGGTGCGCACTGCACTTAAGACCCCT 600
QY 601 CCTATTAACACTTCACGAGAAAGAGACAGTGTCTGACAGAAATCCACCGTGTCTTCT 660
DB 601 CCTATTAACACTTCACGAGAAAGAGACAGTGTCTGACAGAAATCCACCGTGTCTTCT 660
QY 661 GCCCTGCGGAGCTTGCACAAAGGCTTTTGGTAGCTTCGGAACCGTCCGCTCAACGC 720
DB 661 GCCCTGCGGAGCTTGCACAAAGGCTTTTGGTAGCTTCGGAACCGTCCGCTCAACGC 720
QY 721 GGCACGCGCAACGCGCCCTCTGACCAATCTCTCGACGAGCGGAGCAGATCTGACGTT 780
DB 721 GGCACGCGCAACGCGCCCTCTGACCAATCTCTCGACGAGCGGAGCAGATCTGACGTT 780
QY 781 GAGTCGTATTCTCAATGCCCCCTTGAAGGCGGAGCCGCGGATCTCAAGCGAC 840
DB 781 GAGTCGTATTCTCAATGCCCCCTTGAAGGCGGAGCCGCGGATCTCAAGCGAC 840
QY 841 GGGTCTTGGTCTACCGTGAAGAGGCGGCTGAGAGAGTGTCTGCTGTCAATGTC 900
DB 841 GGGTCTTGGTCTACCGTGAAGAGGCGGCTGAGAGAGTGTCTGCTGTCAATGTC 900
QY 901 TACACATGAGACAGCGCTCTGATCAAGCATGCGCTGCGAGAGAAAGCAAGTGCACATC 960
DB 901 TACACATGAGACAGCGCTCTGATCAAGCATGCGCTGCGAGAGAAAGCAAGTGCACATC 960
QY 961 AACGCGTTAGCAACTTTTGTGCGTCAACAACATGATCTACGCTTCAACATCCGCG 1020
DB 961 AACGCGTTAGCAACTTTTGTGCGTCAACAACATGATCTACGCTTCAACATCCGCG 1020
QY 1021 AACGCGACGACGCG 1035
DB 1021 AACGCGACGACGCG 1035
```

RESULT 2
A32185 LOCUS A32185 1035 bp RNA linear PAT 10-DEC-1996

DEFINITION NANBHV PT polyprotein fragment.
ACCESSION A32185
VERSION A32185.1 GI:1926540
KEYWORDS
SOURCE Non-A, non-B hepatitis virus
ORGANISM Non-A, non-B hepatitis virus
REFERENCE Viruses; unclassified viruses.
AUTHORS 1 (bases 1 to 1035)
JOURNAL
FEATURES
source
Patent: FR 2655990-A 4 21-JUN-1991;
Location/Qualifiers
1..1035
/organism="Non-A, non-B hepatitis virus"
/mol_type="unassigned RNA"
/db_xref="taxon:12440"
/clone="JG3"
/clone_1ib="lambda gt11"
<1..>1035
/codon_start=1
/product="partial NANBH-PT polyprotein"
/protein_id="CA02089.1"
/db_xref="GI:1926541"
/translation="TEVDGVRLHRYAACKPLREBYTFOYGLNOYLVSQLPREPER
DVAVLTSLNDPSHTATARKRLAKGSPPELASSASQUSGPKATYITQNPDA
DLIBANLIRHEMGGDITRVESENKVLIDFPDLRAEDREVSVPALIRKSKFP
PAMPAMAPDPNPPLSEWKAPDVPPVHSCPLPPTKTPPIPPPRKRTVLTSTV
SSALAEIATKAFGSSGSAVDSGTATAPDQSDDGAGVSESSMPPLEGPGDP
DLSGWSVTEBEAGBDVVCMSMYTMTGALITPCAEBSLPIINALSNSLRHNMV
YATTSRBSASQR"

ORIGIN

Query Match 100.0%; Score 1035; DB 6; Length 1035;
Best Local Similarity 100.0%; Pred. No. 1.5e-228;
Matches 1035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ACAGAGTGAATGGGGTCCGCTGACAGTACGCTCCGCGTGCAGAACTTCTCTACGG 60
DB 1 ACAGAGTGAATGGGGTCCGCTGACAGTACGCTCCGCGTGCAGAACTTCTCTACGG 60
QY 61 GAGAGGTCAATTCCAGGTCGGGCTCAACCAATACCTGATGGGTGCGAGCTCCATGC 120
DB 61 GAGAGGTCAATTCCAGGTCGGGCTCAACCAATACCTGATGGGTGCGAGCTCCATGC 120
QY 121 GAGCCGGAACCGGATGTAGACAGTGTCTCACTTCCATGCTCAACGACCCCTCCACATGACA 180
DB 121 GAGCCGGAACCGGATGTAGACAGTGTCTCACTTCCATGCTCAACGACCCCTCCACATGACA 180
QY 181 GAGGATGACGGGGAAGTGTCCGCTCCGCGGAGATCTGCGAATTCAGAAATTTCCCA 240
DB 181 GAGGATGACGGGGAAGTGTCCGCTCCGCGGAGATCTGCGAATTCAGAAATTTCCCA 240
QY 241 GCTAGCAGATTGTCTGCGCTTCTCTTGAAGGCGACATATTAATGATTTCCCA 300
DB 241 GCTAGCAGATTGTCTGCGCTTCTCTTGAAGGCGACATATTAATGATTTCCCA 300
QY 301 GACGCTGACCTCATCGAGGCTCAACCTCTGTGTGGCGGATGAGATGGGGCGGACATTACC 360
DB 301 GACGCTGACCTCATCGAGGCTCAACCTCTGTGTGGCGGATGAGATGGGGCGGACATTACC 360
QY 361 CGCGTGAAGTCAAGAAACAAGTGTATCTTGAAGCTTTTGAACCGGCTCCGAGCGGAG 420
DB 361 CGCGTGAAGTCAAGAAACAAGTGTATCTTGAAGCTTTTGAACCGGCTCCGAGCGGAG 420
QY 421 GAGGATGACGGGGAAGTGTCCGCTCCGCGGAGATCTCGGAGAAATTCAGAAATTTCCCA 480
DB 421 GAGGATGACGGGGAAGTGTCCGCTCCGCGGAGATCTCGGAGAAATTCAGAAATTTCCCA 480
QY 481 CCAAGCATGCCCCGATGCGGACGCGCGGATTAACACCTTCGCTGTGAGTCTTGGAAG 540
DB 481 CCAAGCATGCCCCGATGCGGACGCGCGGATTAACACCTTCGCTGTGAGTCTTGGAAG 540
QY 541 GCCCGGACCTACGTCCTCCAGTGTGACATGAGGTGCGCACTGCACTTAAGACCCCT 600
DB 541 GCCCGGACCTACGTCCTCCAGTGTGACATGAGGTGCGCACTGCACTTAAGACCCCT 600
```

Dh 541 GCCCGGAGTACGTCCTCCAGTGTATCATGAGGAGCCCACTTAAGACCCCT 600
Qy 601 CCTATACACTCTCAAGGAGAAAGAGACGTTGTTTGACAGATCAACCGTCTTCT 660
Dh 601 CCTATACACTCTCAAGGAGAAAGAGACGTTGTTTGACAGATCAACCGTCTTCT 660
Qy 661 GCCCTGGAGAGCTTGCCACAAAGGCTTTTGATAGCTCCGAGACCGTGGCGTGCAGC 720
Dh 661 GCCCTGGAGAGCTTGCCACAAAGGCTTTTGATAGCTCCGAGACCGTGGCGTGCAGC 720
Qy 721 GGCACGGCAACCGCCCTCTCTGACCAATCTCCGACGACGGCGAGAGATCTGAGTT 780
Dh 721 GGCACGGCAACCGCCCTCTCTGACCAATCTCCGACGACGGCGAGAGATCTGAGTT 780
Qy 781 GAGTCGTATTCCTTCATGCCCCCTTGAGGGGAGCGCGGGAGACCCCGATCTCAGGAC 840
Dh 781 GAGTCGTATTCCTTCATGCCCCCTTGAGGGGAGCGCGGGAGACCCCGATCTCAGGAC 840
Qy 841 GGGTCTTGGTCTACCGGAGTGAAGAGCGGAGAGAGTGGTCTGCTGCTGATGTC 900
Dh 841 GGGTCTTGGTCTACCGGAGTGAAGAGCGGAGAGAGTGGTCTGCTGCTGATGTC 900
Qy 901 TACACATGACAGAGCGCTGTATCAAGCATGCGCTGCGAGAGAAAGCAAGCTGCCATC 960
Dh 901 TACACATGACAGAGCGCTGTATCAAGCATGCGCTGCGAGAGAAAGCAAGCTGCCATC 960
Qy 961 AAGCGTTGAGCACTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Dh 961 AAGCGTTGAGCACTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Qy 1021 AGGCAAGCCAGCGG 1035
Dh 1021 AGGCAAGCCAGCGG 1035

RESULT 3
ARI44033
LOCUS ARI44033 1035 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 4 from patent US 6210675.
ACCESSION ARI44033
VERSION ARI44033.1 GI:15105900
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1035)
AUTHORS Highfield, P. Edmund., Rodgers, B. Colin., Tedder, R. Seton. and
PT-NANB hepatitis polypeptides
TITLE Patent: US 6210675 A 4 03-APR-2001;
JOURNAL Location/Qualifiers
FEATURES
Source 1. 1035
/organism="unknown"
/mol_type="unassigned DNA"

Query Match 100.0%; Score 1035; DB 6; Length 1035;
Best Local Similarity 100.0%; Pred. No. 1.5e-228;
Matches 1035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAGAGTGTATGGGGTGGCGCTGCAAGGTATGCTCCGGCGTCAAACTCTCTTAAG 60
Dh 1 ACAGAGTGTATGGGGTGGCGCTGCAAGGTATGCTCCGGCGTCAAACTCTCTTAAG 60
Qy 61 GAGAGAGTCAATTCAGAGTGGGCTCAACCAATATCTGTTGGTGGAGCTCCCATGC 120
Dh 61 GAGAGAGTCAATTCAGAGTGGGCTCAACCAATATCTGTTGGTGGAGCTCCCATGC 120
Qy 121 GAGCCGGAACCGGATGTAGAGTGTCTCATCTTCAATGCTCAACGACCCCTCCCATCA 180
Dh 121 GAGCCGGAACCGGATGTAGAGTGTCTCATCTTCAATGCTCAACGACCCCTCCCATCA 180
Qy 181 GCAGAGAGCGCTAAGCGCAAGCTGGCCAGGGGGTCTCCCGCTCTTGCGCACTCTTCA 240

Dh 181 GCAGAGAGCGCTAAGCGCAAGCTGGCCAGGGGGTCTCCCGCTCTTGCGCACTCTTCA 240
Qy 241 GCTAGCAAGTGTCTGAGCCCTTCTCGAAGGCGACATTAATCCAAATATGACTTCCCA 300
Dh 241 GCTAGCAAGTGTCTGAGCCCTTCTCGAAGGCGACATTAATCCAAATATGACTTCCCA 300
Qy 301 GAGCGTGAAGTGTCTGAGCCCTTCTCGAAGGCGACATTAATCCAAATATGACTTCCCA 360
Dh 301 GAGCGTGAAGTGTCTGAGCCCTTCTCGAAGGCGACATTAATCCAAATATGACTTCCCA 360
Qy 361 GCGGTGAAGTCAAGAGAAACAAGTATCTTGAATCTTTTCAACCCGCTCCGAGCGAG 420
Dh 361 GCGGTGAAGTCAAGAGAAACAAGTATCTTGAATCTTTTCAACCCGCTCCGAGCGAG 420
Qy 421 GAGAGTGAAGGAGTGTCTGAGCCCTTCTCGAAGGCGACATTAATCCAAATATGACTTCCCA 480
Dh 421 GAGAGTGAAGGAGTGTCTGAGCCCTTCTCGAAGGCGACATTAATCCAAATATGACTTCCCA 480
Qy 481 CCAAGCATGCGCGATGAGGCGACCGCGGATTAACCCCTCCGCTGCTGAGTCTGGAAG 540
Dh 481 CCAAGCATGCGCGATGAGGCGACCGCGGATTAACCCCTCCGCTGCTGAGTCTGGAAG 540
Qy 541 GCCCGGAGTACGTCCTCCAGTGTATCATGAGTGGCCCACTGCGACCTTAAGACCCCT 600
Dh 541 GCCCGGAGTACGTCCTCCAGTGTATCATGAGTGGCCCACTGCGACCTTAAGACCCCT 600
Qy 601 CCTATACCACTTCACAGGAGAAAGAGACAGTTGTTTGAAGATTCACCGTCTTCT 660
Dh 601 CCTATACCACTTCACAGGAGAAAGAGACAGTTGTTTGAAGATTCACCGTCTTCT 660
Qy 661 GCCCTGGAGAGCTTTCACAAAGGCTTTTGTGATGCTCCGAGACCGTGGCGTGCAGC 720
Dh 661 GCCCTGGAGAGCTTTCACAAAGGCTTTTGTGATGCTCCGAGACCGTGGCGTGCAGC 720
Qy 721 GGCACGGCAACCGCCCTCTCTGACCAATCTCCGACGACGGCGAGAGATCTGAGTT 780
Dh 721 GGCACGGCAACCGCCCTCTCTGACCAATCTCCGACGACGGCGAGAGATCTGAGTT 780
Qy 781 GAGTCGTATTCCTTCATGCCCCCTTGAGGGGAGCGCGGGAGACCCCGATCTCAGGAC 840
Dh 781 GAGTCGTATTCCTTCATGCCCCCTTGAGGGGAGCGCGGGAGACCCCGATCTCAGGAC 840
Qy 841 GGGTCTTGGTCTACCGGAGTGAAGAGCGGAGAGAGTGGTCTGCTGCTGATGTC 900
Dh 841 GGGTCTTGGTCTACCGGAGTGAAGAGCGGAGAGAGTGGTCTGCTGCTGATGTC 900
Qy 901 TACACATGACAGAGCGCTGTATCAAGCATGCGCTGCGAGAGAAAGCAAGCTGCCATC 960
Dh 901 TACACATGACAGAGCGCTGTATCAAGCATGCGCTGCGAGAGAAAGCAAGCTGCCATC 960
Qy 961 AAGCGTTGAGCACTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Dh 961 AAGCGTTGAGCACTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Qy 1021 AGGCAAGCCAGCGG 1035
Dh 1021 AGGCAAGCCAGCGG 1035

RESULT 4
A76573
LOCUS A76573 1788 bp DNA linear PAT 19-OCT-1999
DEFINITION Sequence 5 from Patent WO9317110.
ACCESSION A76573
VERSION A76573.1 GI:608484
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1788)
AUTHORS Rodgers, B.C. and Parker, D.
TITLE A RECOMBINANT HEPATITIS C VIRUS POLYPEPTIDE

JOURNAL Patent: WO 9317110-A 5 02-SEP-1993;

WELLCOME FOUND (GB)

FEATURES

source

CDS

1..1788
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
<1..>1788
/note="Encodes putative PT-NANBH NS5 protein"
/codon_start=1
/product="PT-NANBH coding sequence"
/protein_id="CAB58623.1"
/db_xref="GI:6088485"
/translation="TEVDGRLHRYAPACKPLREBVPQVGLAQVLYGSQLPCEPEP
DVAVLITSLDPSHITAEYAKRLARSGPSLASSASQSGSSKATYITONDFPA
DLIELNLIRHMGDITRVESENKVIILDSFPIRAEDEREVEPAEILAKSKPP
PAMPAPRDYNPRLLESKADYVPVPHGCLPPTKPIPPRRKRTVLTSTV
SSALBARKAFSGSGPSAVDSGTATAPDQSDDGASDVESTSMPELEGFDP
DLSDSWSVSEBAGEDVVCMSYTMVALITPCAESKLPINALNSLRHNMY
YATRSASQROKVTFDRLQILDHYQVLKEMKAKSTVAKLISVEACKLTTPH
SAKSKFGYAKDVRNLSKAINHRSVMDLEDTEPIIDTITAKNEVFCVPERGG
RKPARLIVPDLGVRCESKALYDVVSTLPOAVMGSYGFQYSGPQRYEFLVANKSK
KTPMGFAYTRCPDSTYETENDIRVESTLYCCDLAPEARQAIRSLTERLYIGSLTNS
KQNGCYRRCRASGVLTSCNLTCTYLA"

ORIGIN

Query Match 100.0%; Score 1035; DB 6; Length 1788;
Best Local Similarity 100.0%; Pred. No. 1.4e-228;
Matches 1035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAAAGTGAATGGGGTGGCGGTGACAGAGTACGCTCCGCGGTGCAACCTCTCTACGG 60
DB 1 ACAGAAAGTGAATGGGGTGGCGGTGACAGAGTACGCTCCGCGGTGCAACCTCTCTACGG 60
QY 61 GAGGAGGTACAACTCCAGGTCCGGGTCAACCAATACCTGGTGGGTGAGATCCCATAC 120
DB 61 GAGGAGGTACAACTCCAGGTCCGGGTCAACCAATACCTGGTGGGTGAGATCCCATAC 120
QY 121 GAGCCGGAACCGGATGTACAGTGTCACTTCATCTGACGAGACCCCTCCACATCACA 180
DB 121 GAGCCGGAACCGGATGTACAGTGTCACTTCATCTGACGAGACCCCTCCACATCACA 180
QY 181 GAGAGACGAGCTAAGCGCAGGCTGAGCGAGGGGTCTCCCTCTCTGGCCAGCTTTCA 240
DB 181 GAGAGACGAGCTAAGCGCAGGCTGAGCGAGGGGTCTCCCTCTCTGGCCAGCTTTCA 240
QY 241 GCTAGCAATTTGTCTGGCCCTTCTTGAGAGCGACATATTAATCAACCAATGACTTCCA 300
DB 241 GCTAGCAATTTGTCTGGCCCTTCTTGAGAGCGACATATTAATCAACCAATGACTTCCA 300
QY 301 GAGCGTGAACCTCATCGAGGCCAAACCTCTGTGGCGGATGAGATGGGCGGGGACATTACC 360
DB 301 GAGCGTGAACCTCATCGAGGCCAAACCTCTGTGGCGGATGAGATGGGCGGGGACATTACC 360
QY 361 CGCGTGAAGTCAAGAGAAAGGTAGTATCTTGACTCTTTTGACCCGCTCCGAGCGAG 420
DB 361 CGCGTGAAGTCAAGAGAAAGGTAGTATCTTGACTCTTTTGACCCGCTCCGAGCGAG 420
QY 421 GAGGATGAGCGGGAAGTGTCCGTCGGGGGAGATCTCTCGGGAATCCAGAAATTTCCCA 480
DB 421 GAGGATGAGCGGGAAGTGTCCGTCGGGGGAGATCTCTCGGGAATCCAGAAATTTCCCA 480
QY 481 CCAGCGATCCCGCATGGGACGCGCGGATTAACAACCTCCGCTGTGAGATCTTGGAAG 540
DB 481 CCAGCGATCCCGCATGGGACGCGCGGATTAACAACCTCCGCTGTGAGATCTTGGAAG 540
QY 541 GCCCGGAGTACGTCCTTCAGTGTGATATGGGTGCCCATGCGCATTAAGACCCCT 600
DB 541 GCCCGGAGTACGTCCTTCAGTGTGATATGGGTGCCCATGCGCATTAAGACCCCT 600
QY 601 CCTATACCACTCCACGAGGAAAGAGACAGTTGTTCTGACAGAAATCCACCGTGTCTTCT 660
DB 601 CCTATACCACTCCACGAGGAAAGAGACAGTTGTTCTGACAGAAATCCACCGTGTCTTCT 660

QY 661 GCCCTGGCGAGACTTGCACAAAGGCTTTTGTAGCTCCGAGCCGTCCGCCGTGCACAGC 720
DB 661 GCCCTGGCGAGACTTGCACAAAGGCTTTTGTAGCTCCGAGCCGTCCGCCGTGCACAGC 720
QY 721 GGCACGCAACCGCCCTCTCTGACCAATCTCTCCGACGAGCGCGAGACAGATCTGACGTT 780
DB 721 GGCACGCAACCGCCCTCTCTGACCAATCTCTCCGACGAGCGCGAGACAGATCTGACGTT 780
QY 781 GAGTGTATCTTCATGACGCGCCCTTGAAGGGGAGCGCGGGGAGCCCGGATCTGACGAC 840
DB 781 GAGTGTATCTTCATGACGCGCCCTTGAAGGGGAGCGCGGGGAGCCCGGATCTGACGAC 840
QY 841 GGGTCTTGATCTACCGTGAAGAGAGCCCGGTGAGAGAGTGTCTGCTGCTCATGTC 900
DB 841 GGGTCTTGATCTACCGTGAAGAGAGCCCGGTGAGAGAGTGTCTGCTGCTCATGTC 900
QY 901 TACACATGACAGCGCTCTGATCAAGCCATGCGCTGCGAGAGAAAGCAAGCTGCCATC 960
DB 901 TACACATGACAGCGCTCTGATCAAGCCATGCGCTGCGAGAGAAAGCAAGCTGCCATC 960
QY 961 AACGCTGAGCAACTCTTGTCTGCTGACCAACAATGCTCTACCTACCAATCCCGC 1020
DB 961 AACGCTGAGCAACTCTTGTCTGCTGACCAACAATGCTCTACCTACCAATCCCGC 1020
QY 1021 AGCGCAAGCCAGCG 1035
DB 1021 AGCGCAAGCCAGCG 1035

RESULT 5
A76569
LOCUS A76569 2790 bp DNA linear PAT 19-OCT-1999
DEFINITION Sequence 1 from Patent WO9317110.
ACCESSION A76569
VERSION A76569.1 GI:6088480
KEYWORDS
SOURCE
ORGANISM
unidentified
unclassified.
REFERENCE 1 (bases 1 to 2790)
AUTHORS Rodgers, B.C. and Parker, D.
TITLE A RECOMBINANT HEPATITIS C VIRUS POLYPEPTIDE
JOURNAL Patent: WO 9317110-A 1 02-SEP-1993;
WELLCOME FOUND (GB)
FEATURES
source
location/Qualifiers
1..2790
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
1..>2790
/note="unamed protein product; Fusion of non-structural
and structural protein sequences of PT-NANBH"

CDS
1..2790
/codon_start=1
/protein_id="CAB58621.1"
/db_xref="GI:6088481"
/translation="MPDYSYRPTTGPPPSLSAEFTVDGRLHRYAPACKPLREBY
TFQVGLQVLYGSQLPCEPEPDVAVLITSLDPSHITAEYAKRLARSGPSLASSA
SOLGSPSKATYITONDFPADLIELNLIRHMGDITRVESENKVIILDSFPIRAEDEREVEPAEILAKSKPP
KTPPIPPRRKRTVLTSTVSSALBARKAFSGSGPSAVDSGTATAPDQSDDG
AGSDVESYSWMPLEBEGDPLDSDGSDVTSVEBAGEDVVCMSYTMVALITPCA
ESKLPIINALSNSILRHNMYATRSASQROKVTFDRLQILDHYQVLKEMKAK
ASTYKAKLISVEACKLTTPHSAKSKFGYAKDVRNLSKAINHRSVMDLEDTEPIIDTITAKNEVFCVPERGG
PIDTTMAKNEVFCVPERGGRKPARLIVPDLGVRCESKALYDVVSTLPOAVMGS
YGFQYSPEQRYEFLVANKSKKTPMGFAYTRCPDSTYETENDIRVESTLYCCDLAP
ARQAIRSLTERLYIGPLTNSKQNGCRRRCAGSVLTSCGNLTCTYLAASVAKKK
KATIRANTNLIRPODVRPPGGQIVGVVLLPRGRLGFRATRKTSQSPGRQPI
KARQPEGAMAOPIYMPVLYNREGMAGMLISRGSRPSMGPTDPRRSNLTGVID
TLTGCPADIMGYITPLVAGAPLGAARALAHGVRVLEDDVNTATGNLPGCSFSLFILAL
SCLTIPASAYEVRNVSGLIHYTNDCSNSIYETADIMHTMTPGCVPCYCBENSSRCV
AUIPTLAKDASITPATITRRHVDLVGAAPACSMYVGDLCGSVPFEPQLSAGRYGSF
PGRQREPKTHSLQGNP"

ORIGIN

Query Match 100.0%; Score 1035; DB 6; Length 2790;
 Best Local Similarity 100.0%; Pred. No. 1.4e-228;
 Matches 1035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ACAGAGTGGATGGGGTGGGCTGTCACAGGATCGCTCCGGCGTGCACAAACCTCTCTACGG 60
DB 64 ACAGAGTGGATGGGGTGGGCTGTCACAGGATCGCTCCGGCGTGCACAAACCTCTCTACGG 123
QY 61 GAGGAGTGCATTTCCAGTCTGGGCTCAACCAATACCTGGTGGTGGAGCTCCCATGTC 120
DB 124 GAGGAGTGCATTTCCAGTCTGGGCTCAACCAATACCTGGTGGTGGAGCTCCCATGTC 183
QY 121 GAGCCCGAACCGGATGTAGAGTGTCTCATCTTCCATGCTCACCGAACCCCTCCCATGACA 180
DB 184 GAGCCCGAACCGGATGTAGAGTGTCTCATCTTCCATGCTCACCGAACCCCTCCCATGACA 243
QY 181 GCAGAGAGCGGCTAAGGCGAGGCTGGCCAGGGGGTCTCCCTCTCTGGGCGAGCTCTTCA 240
DB 244 GCAGAGAGCGGCTAAGGCGAGGCTGGCCAGGGGGTCTCCCTCTCTGGGCGAGCTCTTCA 303
QY 241 GCTAGCAGTGTGTCTGGCCCTTCTCTCGAAGGCGCATACATTAACCAAAATGACTTCCCA 300
DB 304 GCTAGCAGTGTGTCTGGCCCTTCTCTCGAAGGCGCATACATTAACCAAAATGACTTCCCA 363
QY 301 GAGCTGACCTCATTCGAGGCCAACCTCTGTGGCGGATGAGTGGCGGGGACATTAAC 360
DB 364 GAGCTGACCTCATTCGAGGCCAACCTCTGTGGCGGATGAGTGGCGGGGACATTAAC 423
QY 361 CGGCTGAGTGCAGAGAACAGATGATGATCTGTGACCTCTTCCAGCCCGCTCCGAGCGGAG 420
DB 424 CGGCTGAGTGCAGAGAACAGATGATGATCTGTGACCTCTTCCAGCCCGCTCCGAGCGGAG 483
QY 421 GAGGATGAGCGGGAGAGTGTCCGTCCGCGCGAGATCTCTGCGAATTCAGAAATTTCCCA 480
DB 484 GAGGATGAGCGGGAGAGTGTCCGTCCGCGCGAGATCTCTGCGAATTCAGAAATTTCCCA 543
QY 481 CCAGCGATGCGCGCATGAGGAGCGCGCGGATTAACAACCTCTCGCTGTGAGTCTTGGAG 540
DB 544 CCAGCGATGCGCGCATGAGGAGCGCGCGGATTAACAACCTCTCGCTGTGAGTCTTGGAG 603
QY 541 GCGCCGGAAGTACGCTCCCTCAAGTGTATCATGGGGGCCCATGCGCACTTACAGACCCCT 600
DB 604 GCGCCGGAAGTACGCTCCCTCAAGTGTATCATGGGGGCCCATGCGCACTTACAGACCCCT 663
QY 601 CCTATACCACTTCACAGGAGAAAGAGAGAGTGTCTGTACAGATCCACCGTGTCTTCT 660
DB 664 CCTATACCACTTCACAGGAGAAAGAGAGAGTGTGTCTGTACAGATCCACCGTGTCTTCT 723
QY 661 GCGCTGCGGAGCTTGGCCACAAAGGCTTTTGTGAGTCTCCGAGCCGTCGCGTGCAGAC 720
DB 724 GCGCTGCGGAGCTTGGCCACAAAGGCTTTTGTGAGTCTCCGAGCCGTCGCGTGCAGAC 783
QY 721 GGCAGCGGCAACCGCCCTCTGACCAATCTCCGACGACGCGGAGAGAGATCTGACGTT 780
DB 784 GGCAGCGGCAACCGCCCTCTGACCAATCTCCGACGACGCGGAGAGAGATCTGACGTT 843
QY 781 GAGTGTATTTCTCTCATGTCGCGCTTGAAGGGGAGCGGGGAGACCCGATCTCAGCGAC 840
DB 844 GAGTGTATTTCTCTCATGTCGCGCTTGAAGGGGAGCGGGGAGACCCGATCTCAGCGAC 903
QY 841 GGGTCTTGGTCTTACCGTGAAGAGAGCGGCTGAGAGAGTCTGTGCTGCTGATGTC 900
DB 904 GGGTCTTGGTCTTACCGTGAAGAGAGCGGCTGAGAGAGTCTGTGCTGCTGATGTC 963
QY 901 TACACATGAGCAGGCGCTGATCAGCGCATGCGCTGCGAGGAGAAAGCAAGTGGCCATC 960
DB 964 TACACATGAGCAGGCGCTTGTATCAGCGCATGCGCTGCGAGGAGAAAGCAAGTGGCCATC 1023
QY 961 AACGCGTTGAGCAACTTGTGCTGCGTCAACCAACATGATCTTACGCTACCATCCGCG 1020
DB 1024 AACGCGTTGAGCAACTTGTGCTGCGTCAACCAACATGATCTTACGCTACCATCCGCG 1083

```

QY 1021 AGGCAAGCCAGCGG 1035
 DB 1084 AGGCAAGCCAGCGG 1098

RESULT 6

A76594 3372 bp DNA linear PAT 19-OCT-1999
 LOCUS Sequence 26 from Patent WO9317110.
 DEFINITION A76594
 ACCESSION A76594
 VERSION A76594.1 GI:6088505
 KEYWORDS

SOURCE unclassified
 ORGANISM unclassified

REFERENCE 1 (bases 1 to 3372)

AUTHORS Rodgers, B.C. and Parker, D.
 TITLE A RECOMBINANT HEPATITIS C VIRUS POLYPEPTIDE
 JOURNAL Patent: WO 93/1710-A 26 02-SEP-1993;
 WELLCOME FOUND (GB)

FEATURES

source

CDS

1..3372
 /note="unnamed protein product; Encodes a fusion protein which can be expressed in insect cells infected with the baculovirus BHC-28"
 /codon_start=1
 /protein_id="CAB58632.1"
 /db_xref="GI:6088506"
 /translation="MPDYSYRPTGPDPSLSAETVEDGVLHRYAPACKPLIREV
 TFOVGNQVLVGSQLPCEPEPVAVLITSLTDPSTHTAETAKRLARGSPSLASSA
 SOLGSSSKATYITONDPDADLIBANLIMRHMGGDITRVESENKVLIDSDPRA
 EDEREVSVPATILRSKKPPPMAMPADPNPPLSEKMAPDYVPVVGCPDLPPT
 KPTPIPPRKXTVUTLSTVSALAELATKARGSGGAGCAVNSTATAPPDSSDDG
 AGSDVDSISMPLEBEPDPPDSDSKSTVSEAEEDVCCSMSTTWGALTTPCA
 ESKLPIINALSNSLHNHNVYATTSRSASQOKKATPRLQLDHYDVLKEMAK
 ASIVKAKLSVEACKLTPPHSKSKFEGYADVRNLSKALINHISVWEDLDEDET
 PIDTTIMAKNEVCVQPERGGRPARLIVPDLGVCEKMAIYDVVSTLPQAMGSS
 YGFOVSPGQVPEFLVNAKSKKTPMGFAVDTECFDSTVENDIRVESITYCCDLPE
 ARATLSLREVLITGPTLNSKQNGVRCASGVLCTSCGNTLTCTKASGILIC
 DECHSTDSISLIGITVLQARTAGKRLVLAATPSPGVTPPHPIEYALSTBEI
 PFYKAIPIETIKGGRHLFCHSKKKCDLALVGLGINAAVYNGILVSVIPASG
 VVVVATDALMTGTDGPDVIDCNCVTOTVDFSLDPTIETTVPODASRSORRG
 RTGRRGRIYRFVAAYKKKKKTKRNTNLRPODVRPPGGQIYGVYLLPRRGRLGV
 RATRKTSSRSOPRGROPPIKAROPEGRAMADPGVPMPIYGNMGMAWMLSPGSR
 PSMGPTDPRRSBNLAKVIDTLTCCPADIMGYIPIVGAFLGAAARLALAGVRLDGV
 NATGMIPGCSFSIFLLALLSCUTTPASAYEVARNVSGIYHVNDCNSNSIYETADMI
 MHPGCVPCVRGNSRCWVALPTPLAADASIPTVITIRRHVDLVGAALFCSAMTVG
 DTCGSVFPEFOLISAGRYGSPFGRORPKTHSLQGNP"

ORIGIN

Query Match 100.0%; Score 1035; DB 6; Length 3372;
 Best Local Similarity 100.0%; Pred. No. 1.3e-228;
 Matches 1035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ACAGAGTGGATGGGGTGGGCTGTCACAGGATCGCTCCGGCGTGCACAAACCTCTCTACGG 60
DB 64 ACAGAGTGGATGGGGTGGGCTGTCACAGGATCGCTCCGGCGTGCACAAACCTCTCTACGG 123
QY 61 GAGGAGTGCATTTCCAGTCTGGGCTCAACCAATACCTGGTGGTGGAGCTCCCATGTC 120
DB 124 GAGGAGTGCATTTCCAGTCTGGGCTCAACCAATACCTGGTGGTGGAGCTCCCATGTC 183
QY 121 GAGCCCGAACCGGATGTAGAGTGTCTCATCTTCCATGCTCACCGAACCCCTCCCATGACA 180
DB 184 GAGCCCGAACCGGATGTAGAGTGTCTCATCTTCCATGCTCACCGAACCCCTCCCATGACA 243
QY 181 GCAGAGAGCGGCTAAGGCGAGGCTGGCCAGGGGGTCTCCCTCTCTGGGCGAGCTCTTCA 240

```

Db	244	GCAGAGACGCGTAAAGCGCAGGCTGCGCCAGGCGGTCTCCCCCTCTTGGCCAGCTCTTCA	303
Qy	241	GCTAGCCAGTTGCTTGGCCCTTCTCGAAGGCGACATATTAACCAAAATGACTTCCCA	300
Db	304	GCTAGCCAGTTGCTTGGCCCTTCTCGAAGGCGACATATTAACCAAAATGACTTCCCA	363
Qy	301	GACGCTGACCTCATCGAGGCGCAACTCTGTGGCGGACATGAGTGGCGGGGACATTAAC	360
Db	364	GACGCTGACCTCATCGAGGCGCAACTCTGTGGCGGACATGAGTGGCGGGGACATTAAC	423
Qy	361	CGCGTGGAGTCAGAGAACAAAGTAGTAATCTTGGACTTTTGGACCCGCTCCGAGCGGAG	420
Db	424	CGCGTGGAGTCAGAGAACAAAGTAGTAATCTTGGACTTTTGGACCCGCTCCGAGCGGAG	483
Qy	421	GAGGATGAGCGGGAAGTGTCCGTCGCGGCGGAGATCTCGCGGAATCAAGAATTCCTCA	480
Db	484	GAGGATGAGCGGGAAGTGTCCGTCGCGGCGGAGATCTCGCGGAATCAAGAATTCCTCA	543
Qy	481	CCAGCGATSCCCGCGATGCGCGCCGCGATTAACAACCTTCGCTGTGAGTCTTGGAG	540
Db	544	CCAGCGATSCCCGCGATGCGCGCCGCGATTAACAACCTTCGCTGTGAGTCTTGGAG	603
Qy	541	GGCCCGGACTAGTCCCTCCAGTGGTACATGGGTGCCCATGCGACCTTAAGACCCCT	600
Db	604	GGCCCGGACTAGTCCCTCCAGTGGTACATGGGTGCCCATGCGACCTTAAGACCCCT	663
Qy	601	CTATATCACTCTCAAGAGAAAGAGAGAGTGTCTTCAAGAAATCCACCGGTCTTCT	660
Db	664	CTATATCACTCTCAAGAGAAAGAGAGAGTGTGTCTTCAAGAAATCCACCGGTCTTCT	723
Qy	661	GGCCTGCGGAGCTTGCCACAAAGGCTTTGTAAGTCTCGGACCGTGGCCGTGACAC	720
Db	724	GGCCTGCGGAGCTTGCCACAAAGGCTTTGTAAGTCTCGGACCGTGGCCGTGACAC	783
Qy	721	GGCAGCGGCAACCGCCCTCTCTCAACCAATCTCTCGAAGAGCGGCGAGACAGATCTACCT	780
Db	784	GGCAGCGGCAACCGCCCTCTCTCAACCAATCTCTCGAAGAGCGGCGAGACAGATCTACCT	843
Qy	781	GAGTCGTATTCCTCCATGCGCCCTCTTGAAGGGGAGCCCGGAGATCTCAGCGAC	840
Db	844	GAGTCGTATTCCTCCATGCGCCCTCTTGAAGGGGAGCCCGGAGATCTCAGCGAC	903
Qy	841	GGGCTCTTGGTCTACCGTGAAGTGAAGAGCCGCTGAGAGACGTCGTCTGCTCGATGTC	900
Db	904	GGGCTCTTGGTCTACCGTGAAGTGAAGAGCCGCTGAGAGACGTCGTCTGCTCGATGTC	963
Qy	901	TACACATGAGACAGGCGCTCTGATCAAGCCATGCGCTGCGGAGAAAGCAAGCTGCCATC	960
Db	964	TACACATGAGACAGGCGCTCTGATCAAGCCATGCGCTGCGGAGAAAGCAAGCTGCCATC	1023
Qy	961	AACGCGTTGAGCAACTTTTGTGCGTCAACAACATGATGATGCTAGCGTCAACATCCGC	1020
Db	1024	AACGCGTTGAGCAACTTTTGTGCGTCAACAACATGATGATGCTAGCGTCAACATCCGC	1083
Qy	1021	AGCGCAAGCCAGCGG 1035	
Db	1084	AGCGCAAGCCAGCGG 1098	
RESULT 7			
A28157			
LOCUS	A28157	3750 bp	DNA
DEFINITION	PT-NANBH mRNA fragment from patent GB2239245.	linear	PAT 04-AUG-1995
ACCESSION	A28157		
VERSION	A28157.1	GI:1248640	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			

FEATURES		Location/Qualifiers
source		1..3750
		/organism="synthetic construct"
		/mol_type="unassigned DNA"
		/db_xref="taxon:32630"
		/note="config formed by cDNA clones from 3' end"
		<1..>3750
CDS		/codon_start=1
		/transl_table=11
		/product="PT-NANBH polypeptide"
		/protein_id="CAA01932.1"
		/db_xref="GI:1248641"
		/translation="MEGVFTGLTVDAHFLSQTRQAGDNPFLVAYQATVCARAQAP
		PSMDQMCLIRLKEPTLRGAPPLRLGAVNEVTLNPIRPFIMACGASALEVST
		FWLGVYLAALAAICLTGSSVIVIGRIISGRPIIVPRLPRLVYQEPDMECSALPLP
		IEQGMQLAEOFKOKALGILQTRATQAEAAAVYSSKRALETTPAKHMNFISGLQV
		AGSLTPENPFIASIMATFAYTLPULTOSTLLNLNIUGVYAAQLAPSAASAVYAG
		IAGAIVSISIGKVLVDILAGVAGVALVAFVMSGEMSTEDLVNLPAILISPGA
		LVGVVCAAILRRHVPGEGAVQMMNRLIAPASGNHVSPTHYVPSDPAARVQIIS
		DLTITQLKRLHWNEDSTPSCGSMRLVDWMICTVLADPTMLQSKLPRLPGVP
		PFSCORGKGVWGRGDMQTTCGGAOITGHVKNKSMRIVGPKTCSNMHGTPLINAY
		TTGCPCTPSPAPNYSBALMRVAAREVYTRVGRDHHVYTSMTDNVKKCCQVPAPEFT
		EVDGRHLRIRYPAKPLRLREVTYQVGLNQYLVSQLPCEPBPVAVITSLTPSHI
		TAETAKRLIARGSPSLASSASQLSAESSVATYITQNDPFDADLIEANLIMRHMGG
		DITRVESENKVIILDSFDPLREBDEREVSVPABILRKSKEPPAMPAPADYNPL
		LESMAAPDYVPVPHGCPLEPTKTPPIPPPRKRTVLTSTVSALAEATKAFGSS
		EPSAVDSGTATAPDOPDDGAGSDVSYSSMPLEBEPDPLSDSMTSVSEAG
		EDVCCSMSTYMTGALIPCAAESKXIPINALSNILRHHMVAATTSRSASOKKY
		TFPRLQIILDDHYQVLEMKRAKASTYVRKALSYEACLTLPHPKAKSFRYGADVRY
		LSSRAINIRSVMEDELDETPTIDITIMANEFVCOPEGRKPAULIVPDLGVR
		VCEKMAIVDVSTLPOAVMSSYGFQYSPGVRVFLVYVAMSKKTPMGFAVDTCFDS
		TVTENDIRVESIYQCCDLPEARQALRSLEIRYVIGGPLTNSGOMNGYRCSAGV
		LRTSCGNLTTCYLASAAKRAKIDQCTMLVCGGLVVICSAAGTQDEAASLRVTEA
		MTRYSAPGDPDPPEYDELITSSSNVSAHDASGRVYVLTFRDP"
ORIGIN		
Query Match	99.1%; Score 1025.4; DB 6; Length 3750;	
Best Local Similarity	99.4%; Pred. No. 2.2e-226;	
Matches 1029; Conservative	0; Mismatches 6; Indels 0; Gaps 0;	
Qy	1	ACAGAAATGAGATGGGGTGGCGCTGCAAGGTACGCTCGGGGTGCAAACTCTCTCTACGG 60
Db	1696	ACAGAAATGAGATGGGGTGGCGCTGCAAGGTACGCTCGGGGTGCAAACTCTCTCTACGG 1755
Qy	61	GAGGAGTCAATTCAGAGTGGGCTCAACCAATTAAGTGGTGGTGGACAGTCCCATGC 120
Db	1756	GAGGAGTCAATTCAGAGTGGGCTCAACCAATTAAGTGGTGGTGGACAGTCCCATGC 1815
Qy	121	GAGCCCGAAGCGATGTAAGAGTGTCACTTCAATGCTCAACCGACCCCTCCCAATCA 180
Db	1816	GAGCCCGAAGCGATGTAAGAGTGTCACTTCAATGCTCAACCGACCCCTCCCAATCA 1875
Qy	181	GCAAGAGAGGCTTAAGCGGAGGTGCGCAGGGGGTCTCCCCCTCTTGGCCAGCTTTCA 240
Db	1876	GCAAGAGAGGCTTAAGCGGAGGTGCGCAGGGGGTCTCCCCCTCTTGGCCAGCTTTCA 1935
Qy	241	GCTAGCAGTTGTCTGCGCCCTTCTCGAAGGCGACATACATTAACCAAAATGACTTCCA 300
Db	1936	GCTAGCAGTTGTCTGCGCCCTTCTCGAAGGCGACATACATTAACCAAAATGACTTCCA 1995
Qy	301	GACGCTGACCTCATCGAGGCGCAACTCTGTGGCGGACATGAGTGGCGGGGACATTACC 360
Db	1996	GACGCTGACCTCATCGAGGCGCAACTCTGTGGCGGACATGAGTGGCGGGGACATTACC 2055
Qy	361	CGCGTGGAGTCAGAGAACAAAGTAGTAATCTTGGACTTTTGGACCCGCTCCGAGCGGAG 420
Db	2056	CGCGTGGAGTCAGAGAACAAAGTAGTAATCTTGGACTTTTGGACCCGCTCCGAGCGGAG 2115
Qy	421	GAGGATGAGCGGGAAGTGTCCGTCGCGGCGGAGATCTCGCGGAATCAAGAATTCCTCA 480
Db	2116	GAGGATGAGCGGGAAGTGTCCGTCGCGGCGGAGATCTCGCGGAATCAAGAATTCCTCA 2175
Qy	481	CCAGCGATSCCCGCGATGCGCGCCGCGATTAACAACCTTCGCTGTGAGTCTTGGAG 540

Db 2176 CAGGATGCGGCGGATGAGGAGCGCCGGGATTAACACCTCCGCTGCTGGATGCTTGGAA 2235
Qy 541 GCCCGGAGTACGTCCTCCATGATGATCATGAGTCCCATCTGCACTTAAGACCCCT 600
Db 2236 GCGCCGAGTACCTCCCTCCAGTGTATCATGGGTGCGCATGCGCACTTAAGACCCCT 2295
Qy 601 CCTATACACCTCCAGGAGAAAGAGACGTTGTTCTGACAAATCCACCGTCTTCT 660
Db 2296 CTTATACACCTCCAGGAGAAAGAGACGTTGTTCTGACAAATCCACCGTCTTCT 2355
Qy 661 GCCCTGCGGAGCTTCCAGCAAAAGCTTTGATAGCTCCGAGCCGTGCGCTGACAG 720
Db 2356 GCGCTGCGGAGCTTCCAGCAAAAGCTTTGATAGCTCCGAGCCGTGCGCTGACAG 2415
Qy 721 GCGAGCGCAACCGCCCTCTCTGACCAATCTCCGAGCGCGAGAGATTTGACCT 780
Db 2416 GCGAGCGCAACCGCCCTCTCTGACCAACCTCCGAGCGCGAGAGATTTGACCT 2475
Qy 781 GAGTGTATTTCTCATGCGCCCTTTGAGGGGAGCGCGGAGACCCGATCTCAGGAC 840
Db 2476 GAGTGTATTTCTCATGCGCCCTTTGAGGGGAGCGCGGAGACCCGATCTCAGGAC 2535
Qy 841 GGGTCTTGTCTACCGTGAATGAGAGCGCGTGAAGACGTCGTCGCTGATGTC 900
Db 2536 GGGTCTTGTCTACCGTGAATGAGAGCGCGTGAAGACGTCGTCGCTGATGTC 2595
Qy 901 TACACATGACAGCGCTGATGACCGCATGCGCTGCGAGAGAAAGCAAGTCCCATC 960
Db 2596 TACACATGACAGCGCTGATGACCGCATGCGCTGCGAGAGAAAGCAAGTCCCATC 2655
Qy 961 AAGCGCTTGAAGCACTTTGCTGCGTCAACCAATGCTTACGATCAATCCCGC 1020
Db 2656 AAGCGCTTGAAGCACTTTGCTGCGTCAACCAATGCTTACGATCAATCCCGC 2715
Qy 1021 AGCGCAAGCCAGCGG 1035
Db 2716 AGCGCAAGCCAGCGG 2730

RESULT 8
A32203 3750 bp DNA linear PAT 10-DEC-1996
LOCUS NANBHV PT polypeptide fragment.
DEFINITION A32203
ACCESSION A32203.1 GI:1926564
VERSION A32203.1
KEYWORDS Non-A, non-B hepatitis virus
SOURCE Non-A, non-B hepatitis virus
ORGANISM Viruses; unclassified viruses.
REFERENCE 1 (bases 1 to 3750)
AUTHORS Patent: FR 2655990-A 22 21-JUN-1991;
JOURNAL Location/Qualifiers
FEATURES
source 1..3750
/organism="Non-A, non-B hepatitis virus"
/mol_type="unassigned DNA"
/db_xref="taxon:12440"
CDS
1..>3750
/codon_start=1
/product="partial NANBH-PT polypeptide"
/db_xref="GI:1926565"
/translation="MEGVFTGLTHVDHFLSQTKQADNFPYLVAVQATVCARQAQAP
PSWDMWKLRLKLTGRPTPLRYLGAIVONVETLHPITKRIACMSADLEVVT
WLVGVGLAALAYCLTGSVIVGRILISGRPAIVDEVLVQEDMEBESHLY
IRGQMLAOFKOKALGLQATKOAEAPAVVESRRALETFMAKMNPIGILY
AGLSTLPNPAIATLMAFTASVTSPLTOSTLILNIGVNPNSLPBAASA
FVAGTAGAVSISGLKTVLDIILAGYAGVAGALVAFKMSGEMPTEDLVNLPALISGA
LVAVGVCALIRRHVPGEGAVQVMNRLIAFASRANHVSPTHVPSDAAARQTLLS
DLITLQRLHQMINEPSCSGMLNDWDMICTVLADFTWLSKLPPLTQLS
PFSQGYGVNMGDGIQOTSCGAQITGVNMGSMRIAGPTCSNMHGTPIVAV
TTGPTCPSPAPNYSRALMRLVAAEYEVETRVGD FHVVTSMITDNVACPGQVAPBEFT

ORIGIN
Query Match 99.1%; Score 1025.4; DB 6; Length 3750;
Best Local Similarity 99.4%; Pred. No. 2,2e-226;
Matches 1029; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
1 ACAGAGTGAATGAGGCTGCGGCTGACAGATGAGCTCCGAGCGCAACCTCTCTACGG 60
Db 1696 ACAGAGTGAATGAGGCTGCGGCTGACAGATGAGCTCCGAGCGCAACCTCTCTACGG 1755
Qy 61 GAGAGGTCAATTCAGGTGCGGCTCAACCAATACCTGATGAGTCCGATGC 120
Db 1756 GAGAGGTCAATTCAGGTGCGGCTCAACCAATACCTGATGAGTCCGATGC 1815
Qy 121 GAGCCGGAACCGGATGATGAGTGCCTCACTTCACTGATCCGACCCCTCCACATCA 180
Db 1816 GAGCCGGAACCGGATGATGAGTGCCTCACTTCACTGATCCGACCCCTCCACATCA 1875
Qy 181 GAGAGAGCGGCTAAGGCGAGGCTGCGGAGGAGTCCGCCCTCTTGGCAGCTTTCA 240
Db 1876 GAGAGAGCGGCTAAGGCGAGGCTGCGGAGGAGTCCGCCCTCTTGGCAGCTTTCA 1935
Qy 241 GCTAGCAGTTGTCTGCGCTTCTCTGAGAGGAGCATATTAACCAAAATGA 300
Db 1936 GCTAGCAGTTGTCTGCGCTTCTCTGAGAGGAGCATATTAACCAAAATGA 1995
Qy 301 GACGCTGACCTCATGAGGCGCACTCTCTGTGGCGGATGAGTGGCGGAGCATTAAC 360
Db 1996 GACGCTGACCTCATGAGGCGCACTCTCTGTGGCGGATGAGTGGCGGAGCATTAAC 2055
Qy 361 CGCGTGAAGTCAAGAGCAAGGATGATCTCTGACCTTTGACCGGCTCCGAGCGAG 420
Db 2056 CGCGTGAAGTCAAGAGCAAGGATGATCTCTGACCTTTGACCGGCTCCGAGCGAG 2115
Qy 421 GAGGATGAGCGGAGATGTCCTGCTCCGCGGAGATCTCGGAATCAAGAAATCCCA 480
Db 2116 GAGGATGAGCGGAGATGTCCTGCTCCGCGGAGATCTCGGAATCAAGAAATCCCA 2175
Qy 481 CCAAGCATGCGCGCATGAGGCGCGCGGATTTACACCTCTCTGAGTCTGGAAG 540
Db 2176 CCAAGCATGCGCGCATGAGGCGCGCGGATTTACACCTCTCTGAGTCTGGAAG 2235
Qy 541 GCCCGGAGTACGTCCTCCAGTGTATGAGTGGTCCCATCTGACCTTAAGACCCCT 600
Db 2236 GCCCGGAGTACGTCCTCCAGTGTATGAGTGGTCCCATCTGACCTTAAGACCCCT 2295
Qy 601 CCTATACACCTCCAGGAGAAAGAGACGTTGTTCTGACAAATCCACCGTCTTCT 660
Db 2296 CCTATACACCTCCAGGAGAAAGAGACGTTGTTCTGACAAATCCACCGTCTTCT 2355
Qy 661 GCCCTGCGGAGCTTCCAGCAAAAGCTTTGATAGCTCCGAGCCGTGCGCTGACAG 720
Db 2356 GCCCTGCGGAGCTTCCAGCAAAAGCTTTGATAGCTCCGAGCCGTGCGCTGACAG 2415
Qy 721 GCGAGCGCAACCGCCCTCTCTGACCAATCTCCGAGCGCGAGAGATTTGACCT 780
Db 2416 GCGAGCGCAACCGCCCTCTCTGACCAACCTCCGAGCGCGAGAGATTTGACCT 2475
Qy 781 GAGTGTATTTCTCATGCGCCCTTTGAGGGGAGCGCGGAGATTTGACCT 840
Db 2476 GAGTGTATTTCTCATGCGCCCTTTGAGGGGAGCGCGGAGATTTGACCT 2535

QY 841 GGGCTTGGTCTACCGTGAAGAGAGCGCGGTGAAGACGTCTGTCTGATGTC 900
|
|
|
Db 2536 GGGCTTGGTCTACCGTGAAGAGAGCGCGGTGAAGACGTCTGTCTGATGTC 2595
|
|
|
QY 901 TACACATGACAGCGCGCTCTGATACGCCATGCGCTCGGAGAAAGAACTGCCATC 960
|
|
|
|
Db 2596 TACACATGACAGCGCGCTCTGATACGCCATGCGCTCGGAGAAAGAACTGCCATC 2655
|
|
|
|
QY 961 AACGGCTTGAACAACCTTGTGCGGTGACCAACAATGCTGACGCTACCAATCCGCG 1020
|
|
|
|
Db 2656 AACGGCTTGAACAACCTTGTGCGGTGACCAACAATGCTGACGCTACCAATCCGCG 2715
|
|
|
|
QY 1021 AGCGCAAGCCAGCGG 1035
|
|
|
|
Db 2716 AGCGCAAGCCAGCGG 2730
|
|
|
|
RESULT 9
ARI44051 3750 bp DNA linear PAT 08-AUG-2001
LOCUS Sequence 22 from patent US 6210675.
ACCESSION ARI44051
VERSION ARI44051.1 GI:15105918
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3750)
AUTHORS Highfield,P.Edmund., Rodgers,B.Colin., Tedder,R.Seton. and
Barbara,J.Anthony.James.
TITLE PT-NANB hepatitis polyprotein
JOURNAL Patent: US 6210675-A 22 03-APR-2001;
FEATURES
source 1..3750
/organism="unknown"
/mol_type="unassigned DNA"

Query Match 99.1%; Score 1025.4; DB 6; Length 3750;
Best Local Similarity 99.4%; Pred. No. 2.2e-226;
Matches 1029; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 ACAGAGTGAATGGGTGCGGCTGCACAGTACGCTCCGCGGTGCAACCTCTCTACCG 60
|
|
|
Db 1696 ACAGAGTGAATGGGTGCGGCTGCACAGTACGCTCCGCGGTGCAACCTCTCTACCG 1755
|
|
|
QY 61 GAGGAGTCAATTCAGGTCCGGCTCAACCAATACCTGTTGGTTCGACGCTCCATC 120
|
|
|
|
Db 1756 GAGGAGTCAATTCAGGTCCGGCTCAACCAATACCTGTTGGTTCGACGCTCCATC 1815
|
|
|
|
QY 121 GAGCCGGAACCGGATGACAGTCACTTCATGCTCAACGACCCCTCCCAATCA 180
|
|
|
|
Db 1816 GAGCCGGAACCGGATGACAGTCACTTCATGCTCAACGACCCCTCCCAATCA 1875
|
|
|
|
QY 181 GAGAGACCGCTAAGCGAGGCTGCGCAAGGAGTCTCCCTCTCTTGGCCAGCTCTTA 240
|
|
|
|
Db 1876 GAGAGACCGCTAAGCGAGGCTGCGCAAGGAGTCTCCCTCTCTTGGCCAGCTCTTA 1935
|
|
|
|
QY 241 GCTAGCACTTGTCTGCGCTTCTCTGAAGGCAATACATTAACCAAAATGACTTCCA 300
|
|
|
|
Db 1936 GCTAGCACTTGTCTGCGCTTCTCTGAAGGCAATACATTAACCAAAATGACTTCCA 1995
|
|
|
|
QY 301 GAGCGTCACTCATGAGGCAACCTCCGTGGCGGCAAGATGGCGGGGACATTAAC 360
|
|
|
|
Db 1996 GAGCGTCACTCATGAGGCAACCTCCGTGGCGGCAAGATGGCGGGGACATTAAC 2055
|
|
|
|
QY 361 CGCGTGAATCAGAGAAACAAGTATGATCTGACTCTTTCGACCCGCTCCGAGCGAG 420
|
|
|
|
Db 2056 CGCGTGAATCAGAGAAACAAGTATGATCTGACTCTTTCGACCCGCTCCGAGCGAG 2115
|
|
|
|
QY 421 GAGGATGACGGGAAAGTGTCCGTCGCGGAGAGATCTTCGGAATCCAAGAATTTCCA 480
|
|
|
|

Db 2116 GAGGATGACGGGAAAGTGTCCGTCGCCGCGGAGATCTTCGGAATCCAAGAATTTCCA 2175
|
|
|
QY 481 CCAGGATGCCCGCATGGGACCGCCGGATTACAACCTCCGCTGTGAGGCTCGAAG 540
|
|
|
|
Db 2176 CCAGGATGCCCGCATGGGACCGCCGGATTACAACCTCCGCTGTGAGGCTCGAAG 2235
|
|
|
|
QY 541 GCCCGGACCTACGTCCTCCAGTGTATCATGGGTGCCCACTTAAAGACCCCT 600
|
|
|
|
Db 2236 GCCCGGACCTACGTCCTCCAGTGTATCATGGGTGCCCACTTAAAGACCCCT 2295
|
|
|
|
QY 601 CCTATACCACTCCACGAGAAAGACAGTTGTTCTGAACAATCCACCTGCTTCT 660
|
|
|
|
Db 2296 CCTATACCACTCCACGAGAAAGACAGTTGTTCTGAACAATCCACCTGCTTCT 2355
|
|
|
|
QY 661 GCCCTGGGAGCTTGGCAAAAGCTTTTGTGATGCTCCGAGCCGTGGCCGTGAGACG 720
|
|
|
|
Db 2356 GCCCTGGGAGCTTGGCAAAAGCTTTTGTGATGCTCCGAGCCGTGGCCGTGAGACG 2415
|
|
|
|
QY 721 GGCACGGCAACCGCCCTCTGACCAATCTCCGACGCGGAGACAGATCTGACGTT 780
|
|
|
|
Db 2416 GGCACGGCAACCGCCCTCTGACCAATCTCCGACGCGGAGACAGATCTGACGTT 2475
|
|
|
|
QY 781 GAGTCTATTTCTCTCATGCCCCCTTTGAGGGGAGACCGGAGACCCGATCTACGAC 840
|
|
|
|
Db 2476 GAGTCTATTTCTCTCATGCCCCCTTTGAGGGGAGACCGGAGACCCGATCTACGAC 2535
|
|
|
|
QY 841 GGGCTTGGTCTACCGTGAAGAGAGCGGTGAGAGCGTCTGCTGCTGATGTC 900
|
|
|
|
Db 2536 GGGCTTGGTCTACCGTGAAGAGAGCGGTGAGAGCGTCTGCTGCTGATGTC 2595
|
|
|
|
QY 901 TACACATGACAGCGGCTCTGATACGCCATGCGCTCGGAGAAAGCACTGCCATC 960
|
|
|
|
Db 2596 TACACATGACAGCGGCTCTGATACGCCATGCGCTCGGAGAAAGCACTGCCATC 2655
|
|
|
|
QY 961 AACGGTTGAGCAACTCTTGTCTGCTGCTACCAACAATGCTTACCAATCCGCG 1020
|
|
|
|
Db 2656 AACGGTTGAGCAACTCTTGTCTGCTGCTACCAACAATGCTTACCAATCCGCG 2715
|
|
|
|
QY 1021 AGCGCAAGCCAGCGG 1035
|
|
|
|
Db 2716 AGCGCAAGCCAGCGG 2730
|
|
|
|
RESULT 10
A76577 7065 bp DNA linear PAT 19-OCT-1999
LOCUS Sequence 9 from Patent WO9317110.
ACCESSION A76577
VERSION A76577.1 GI:6088488
KEYWORDS
SOURCE unidentified
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 7065)
AUTHORS Rodgers,B.C. and Parker,D.
TITLE A RECOMBINANT HEPATITIS C VIRUS POLYPEPTIDE
JOURNAL Patent: WO 9317110-A 9 02-SEP-1993;
FEATURES
source 1..7065
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
<1..7065
/note="Bncodes putative PT-NANBH non-structural proteins"
/codon_start=1
/product="PT-NANBH coding sequence"
/protein_id="CAB58625.1"
/db_xref="gi:6088489"
/translation="SCGAVPIGLMPLTISPYRVPLAKLIMTQYITRAEHLQW
VPLNVRGRDAITLACVHBEIFPDISKLILILGMLMLQGITRVYFBAQGL
IRACMLVRKAVGHHVQMALMKLAALGTYYDHLTFLRDAHAGLDAVAEPPVR
SDMETKVIWTGDTAAAGDITLIPVBARBDEIFLDPADSLBGOGRLLAPITAYSO

OTRGLGCIITSLTGRDKNOVEGEVOVNSTATOSPLATCVNGCMTVYHGAAGSKTLG
PGSPVOMTYNDODLVGMPAPRGASRLPCTCGSSDILYTHRHAVIYPRRRDGRG
SILSPRIYVINGSSGGPILCPBGHVGIFRAVCSCTRGVAKANDPVPBSEMETTVASG
VTFDNSPPRAVPOSFVVAHLHAPVSGSKSTRVPAAYAAQCYKVLVNPSPVMAVCLGGA
YMSKAGVDPNIRKSVGRTITVGAIPITYSTYKFLADGGCGGAYDIIICECHSTST
SILGICVTDVDOAETAGARLVLATATPPGVSPHPNIEBVLSTNNGIPEYKALPI
ERTKGRHLIFCHSKKCDLAKLVGLGNVAVAYTIPASGDVAVATATL
MNGFTGDPVSIDONTCTVOTVDPISLDPTEITETTVPODVAVSORBRGTGGRGI
YRFVAGERPSGMFDSVCECYDAGCANTELTPAETSVRLAYLNTPGIPVODHLE
FNESTFGLIHDVDAFLSOTKQDNFPIVAYATQACAAQAPSPSWDMWCLIRL
KPTLRGPTLVRGLAVONEVTLTHPIKIMACMSADIEVSTVWLVGVVLAALAA
YCLTGSVVIYGR11ILSGRPAIVPDRLEVYQEFDEMEBEGSHLYIEQGLAEQKQ
KALGLTOTAKOEAAPVYESKRALETPMKMMNFISIGIYLAGLSTLPGNPAIA
SILMAFTASVTSPLTOSTILLNIGMGVAAQLAPSAASFGVAGAGVCAITIGK
VLVDILAGVAGAGVAGVAFKMSGEMPTEDLVNLLPALSFGALVGVCAITIRK
HVGPGGAVQMMNRLIAFASRGNHSPHIVPESDAARVTLQSLDTYQILKRLHO
WINECSTPSCSGMIRVDVMICTVLADPFTWLOSILPLPVPFSCORGKXKAVY
GDGIMQTCSCGQITGVHNGSMRIVGPTCSNMHMGPPINAITYGCTPSPBANY
SRALMVAABEYEVTRVGDPHYVTSMITDNVACPOVPAPEFTFVDGRLHRYVA
CKPLABEYTFVYGLNOYLVGSQLPCEPBDVAVLSTMLTPBHTAETRAKRLAGS
PSRLASSASQLSAPSSKATYITONDPPBDLIEANILRHHEMGITTRYESNKVIT
LUSFEDPLRAEBDEREVSVAEILIRKSKPEPPAMPAPARPDVNPILLESWAPVYVP
VHGCPLEPTKTPPIPPRRKRTVLTSTVSALAEIATAFSGSPSADVSGTAPAY
POPSDGGAGSIVESYSSMPLEGGPBDLSGMSVTSBAGSDVVCSSMYMT
GALITPCAABESKLPIINALSNSILRHHNVYATTSRSASOROKKTPEDRLQILDDHYO
DYLKEMKAKASTYKALKLSVEBAKLTPHSAKSKRGYAKDRNLSSKAINIRSVY
EDLEBETPITDTTINAKNEVFCVQPERGKRPRLVFPDLCVRYCEKALYDVST
LPOAVNGSSYFQYSPQORVEFLVNAKSKRTMPRPAVDRDSCVTEVNDLREEVSK
AKLSVEBAKLTTPHSAKSKFGYKADVNLSKALINIRSWEDLLEBETPITDTT
INAKNEVFCVQPERGKRPRLVFPDLCVRYCEKALYDVSTLPOAVNGSSYGOY
SPQORVEFLVNAKSKRTMPRPAVDRDSCVTEVNDLREEVSKYOCCEPARARAI
RSLTERLYIGPLNPKNSKNGONCYRCRAGVLTSCGNTITCYKASACRAKALODC
THLYCDDLVYICESAGTQDASLAVTFEAMTRYSAPREDPQPEYDELTJSCSN
VSVADHDSGRVYLTLDPTTPLARAMEARTPVSNGIINMAIPTAIIWAAILMT
HEFESILAOBOLKALDCOYGAYSIEPDLPOIIOURLGASISHSPEIRNV
ASCLRLGIVPLRMRHRAVAVAKLSOGGRAICRGYLFMAVATKLTPIPAAS
RLDLSGMFVAGYGGDIIYHSVBARPRMFWCILLISVGVILLENR"
7063 7065
/function="Stop codon"

ORIGIN

Query Match 99.1%; Score 1025.4; DB 6; Length 7065;
Best Local Similarity 99.4%; Pred. No. 2.1e-226;
Matches 1029; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 ACAGAGTGGATGGGTCGGCTGACAGGTAGCTTCGGCGTGAACCTCTCTTAAG 60
DB 3919 ACAGAGTGGATGGGTCGGCTGACAGGTAGCTTCGGCGTGAACCTCTCTTAAG 3978
QY 61 GAGGAGTCACTTCCAGGTGGGCTCAACCACTACTGTTGGGTGGAGCTCCCAATGC 120
DB 3979 GAGGAGTCACTTCCAGGTGGGCTCAACCACTACTGTTGGGTGGAGCTCCCAATGC 4038
QY 121 GAGCCCGAACCAGATGAGAGTGTCACTTCATGCTCCAGACCCCTCCCATCATCA 180
DB 4039 GAGCCCGAACCAGATGAGAGTGTCACTTCATGCTCCAGACCCCTCCCATCATCA 4098
QY 181 GCAGAGACGGCTAAGCGCAGAGCTGGCCAGGGGGTCTCCGCCCTCTTGGCCAGCTCTTCA 240
DB 4099 GCAGAGACGGCTAAGCGCAGAGCTGGCCAGGGGGTCTCCGCCCTCTTGGCCAGCTCTTCA 4158
QY 241 GCTAGCAGATTGTCTGGCCCTTCTCTGGAAGCGACATATACCAACCAATGACTTCCCA 300
DB 4159 GCTAGCAGATTGTCTGGCCCTTCTCTGGAAGCGACATATACCAACCAATGACTTCCCA 4218
QY 301 GACGCTGACCTCATCGAGGCCAATCTCTTGGCGGCGATAGATGGCGGGGACATTACC 360
DB 4219 GACGCTGACCTCATCGAGGCCAATCTCTTGGCGGCGATAGATGGCGGGGACATTACC 4278
QY 361 CGGCTGAGTCAAGAAACAAGTAGTAACTCTGAACTCTTTCGACCCGCTCGAGCGGAG 420
DB 4279 CGGCTGAGTCAAGAAACAAGTAGTAACTCTGAACTCTTTCGACCCGCTCGAGCGGAG 4338

QY 421 GAGATGAGCGGGAAGTGTCCGTCGGCGGAGATCTCGGAAATCCAGAAATTTCCCA 480
DB 4339 GAGATGAGCGGGAAGTGTCCGTCGGCGGAGATCTCGGAAATCCAGAAATTTCCCA 4398
QY 481 CCAGCATGCGCGATGGGACGCGCCGGATTACACCTTCGCTGTGAGTCTGGAAG 540
DB 4399 CCAGCATGCGCGATGGGACGCGCCGGATTACACCTTCGCTGTGAGTCTGGAAG 4458
QY 541 GCCCGGATGTAAGTCCCTCCAGTGTGATCATGGGTGCCCATCTGCACCTTAAGACCCCT 600
DB 4459 GCCCGGATGTAAGTCCCTCCAGTGTGATCATGGGTGCCCATCTGCACCTTAAGACCCCT 4518
QY 601 CCTATACACCTTCCAGGAGAAAGAGACAGTTGTTCTGACAGATTCACCCGCTCTTCT 660
DB 4519 CCTATACACCTTCCAGGAGAAAGAGACAGTTGTTCTGACAGATTCACCCGCTCTTCT 4578
QY 661 GCCCTGCGGAGCTTGCACAAAGGCTTTGTAGCTCCGACCGTCCGCGTGAACAG 720
DB 4579 GCCCTGCGGAGCTTGCACAAAGGCTTTGTAGCTCCGACCGTCCGCGTGAACAG 4638
QY 721 GGCACGGCAACCGCCCTCTCTGACCAATCTCCAGACAGCGGAGACAGATCTGACGTT 780
DB 4639 GGCACGGCAACCGCCCTCTCTGACCAATCTCCAGACAGCGGAGACAGATCTGACGTT 4698
QY 781 GAGTGTATTCTCTCAATGCCCCCTTGAAGGGGAGCCGGGGGACCCCGATCTCAGCGAC 840
DB 4699 GAGTGTATTCTCTCAATGCCCCCTTGAAGGGGAGCCGGGGGACCCCGATCTCAGCGAC 4758
QY 841 GGGTCTTGTGTTACCGTGAAGTGAAGAGCGCGGTGAGAGCGTCTGCTGCTGATGTC 900
DB 4759 GGGTCTTGTGTTACCGTGAAGTGAAGAGCGCGGTGAGAGCGTCTGCTGCTGATGTC 4818
QY 901 TACACATGACACGCGCTCTGATCAGCCATGCGCTGCGGAGAAAGCAAGCTGCCATC 960
DB 4819 TACACATGACACGCGCTCTGATCAGCCATGCGCTGCGGAGAAAGCAAGCTGCCATC 4878
QY 961 AACGCTTGAAGCACTTGTGCTGCGTCAACCAATGATGATCTGACATCCCGC 1020
DB 4879 AACGCTTGAAGCACTTGTGCTGCGTCAACCAATGATGATCTGACATCCCGC 4938
QY 1021 AGCGCAAGCCAGCGG 1035
DB 4939 AGCGCAAGCCAGCGG 4953

RESULT 11
AF313916
LOCUS AF313916 9359 bp mRNA linear VRL 01-JAN-2002
DEFINITION Hepatitis C virus polyprotein mRNA, partial cds.
ACCESSION AF313916
VERSION AF313916.1 GI:18027684
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
Virus; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE 1 (bases 1 to 9359)
AUTHORS Fanning,J., Itakura,J., Nagayama,K. and Enomoto,N.
TITLES Characteristics of Hepatitis C viral genome associated with disease
progression in a homogeneous patient population
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 9359)
AUTHORS Fanning,J., Itakura,J., Nagayama,K. and Enomoto,N.
TITLES Direct Substitution
JOURNAL Submitted (17-OCT-2000) Medicine, National University of Ireland,
Cork, Hepatitis C Unit, Clinical Sciences Building, Cork University
Hospital, Cork, Ireland
FEATURES
source
1..9359
/organism="Hepatitis C virus"
/mol_type="mRNA"
/db_xref="taxon:11103"
/note="Isolated from source plasma identified as the

single source of the iatrogenic infection of Rheus negative women in Ireland from May 1977-November 1978; exposure was through HCV 1b contaminated anti-D immunoglobulin"

5'UTR
CDS

330...>9359
/codon_start=1
/product="polypotein"
/protein_id="AA155821.1"
/db_xref="GI:16027685"

/translation="MSTNPKPQRTKRNTRRPQVKKPFGGQIVGVYLLPRRPRRL
GRATIRKTSERQPRGRQPIPKARPERTAQOPGYEMPLVGNKGAGMGLVLEPRG
SRPSGPTDPRRRSRNLKVIDITLTCGFDLMGYIPLVAGAGARLALHGVRLIED
GNVATGALPGCSFSIPLIALISCLTIPASALEVNSGVYVNDLSAIVYAAD
MIMHTGCVPCRENNSSRCVVALPPTLAARNSIIPITIRHVTLVLTGASIVYAEAD
VGDLCGVFLVSDLTFTSPRHEIWDQNCISYPGHVSHHMAAMMMNSPTTLVY
SGLARIPOAVDUMAGAHVGLAGLAIYSMAKNKXVLVLLPFGVDGETTGTGAQ
AHTRGATLTFSPSQNTQLINTGSMHINTALNCDSLNTGFLALFVAHRNAS
GCPERMASCRPIDKPAQGWPIAYAKPLSDPKCMHAPQPCIVAPASQVCGVYC
FPPSVVVGTDRPGVPTYSWGENEDVLLNTRPBRGNMFGCTMMNGTPTKTCG
PCDLCGVNDTLICPTDCFRKHPKATYKCGSGPMLPRCWVDYRLMHYPCVNF
TIFKRMVYGVGEHRLNACMWTGRBCDLEDRSSELPLSTTEMQVPLCSPTTL
PALSTGLIHLHRIYDVQYLYIGISAVSYAIKREYVLLFTFLADARYCAGLMMLL
IVQAEALENLVINAASVAGTHGISLPLFCAMAYIKGRVLPGAAYAFGVWPLLL
LILALPRAVYAMDREMAASCQAVFGLALTLSPHYKVFARLIMVLYQVITRAEAL
LQVWIPPLAVRGDPAIILMCAIHBEIIFATKILILGLPVLQAGITRPFYFR
AHGLIRACVIREVAGHYOMALMRALATGYVYDHLTIPDMAHSELRDAVAVE
PVFSDMETKIIITWGDPTAAGDIIIGLPVSRARGBEILIGADSLBOGMBLADIT
AYAOOTRGLGCTIVSLTRDKNQVGEVQVYSTAOSTALCNVGCCTVYHAGTK
TLAQQGPIITQMTVNDODVGMALPGRASLPCTCGSDILVTRHADVLVARRG
DSRGLSPRPVSYLKSGSGPLCPGHAAGI FRAAVCTRVAAAVFVPEVSEMTT
MSPVFTDNSSPPAVPOTFOVAMHAPTSKSGSTVPAAYAGQKAVLVNPSVATL
GFGAAMSKAHGVDPNIRGTARTITGAATYTSYTGKFLGSGAGVADIIICDSCS
TDSITLIGTGLVDAETGARTLVLAATPBGSVTVHPNIEVALNIGIIPYCHG
AIPITIGGRSLIECHSKKCDLAAKLSGLANAVAYRDLVSVIPSGDVVVA
TDALMTGTGDFSDVDCNTCVTQVDSLDPTFTILETIVQDAAVPTSGDVVVA
RTGIYRPTVPGERPSGEMDSVLCBYCAGCAMYELTPAETSVLRALANTPGLPVCQ
DHLFEWEGFTLTHIDAFLSQTKOAGNPYLAQVATVCAQAAQAPPSVDWMKC
LIRLKEPTLHPTLYRLGAVONEVILHPTIKFIMACASALEVATSTWVVGSLA
ALAAVCLTTSVSIVIRIILSGKPAIVPBEVLYOOPMEMEACHLPIITRGMOLAE
OPKOKAGLLOTRATQABAAAVYBSKRALESFAKMMFISGQIYLAISTIPGN
PAIASMAFTASITPLOTHTLNFILIGVAVAQALAPPASAVAGAGAGAAVSTIPGN
IGKXVLVDILAGVAGVAGALVAFVMSGEMSTEDLVNLLPAILSPALVGVVCA
ILRHVYGBEGAVQMMNRLIAPASRGNYSPHYVESPASVLIQIISLITLTLK
RLHOWINEDCSPSGSMRLDWMICVLTLDPTKMSKILPRLPGPSPSCOGYK
GVMRGDMQCTPCGAOITGHVKNKSRIVGRTKNTMGTPTIPNAYTGPCTPSP
APNYSRALMRVTAEEYIELVYQGDPHYVYGMTDVKPCQVPAPEPTFEDVCYLR
YAPCKPLAREVITFOVLANQYLVESQYLCPEPVAVYVLSMLTOSHTLAARLRL
AKGSPSLASSSASQLSAPLSKATCTTRHSDPADLIEANILMRQEMGNTIRYSEN
KVVLIDSEPLAEEDEREVSYPALILRSRKFPRAMPIMADPTNFPSPLESKDPDY
VPPVHAGCPLEPTKAPIPEPRKRITVLTESVSALAEATKTFNPSSESADVSGT
ATAPDPOSDDDAGSDVESYSMPLSEBEPDPLISGSPVTESEASBPVCCSMS
YTMGALITPCSAEESKLPINALSMSILRHHMYATTSRSRSQOKQVCTPRLOYLD
DHYRVLKMKRKAATVAKLLPVEBACLTTPHSAKSFEGAGADVKNLSKAVNHI
LSVMKDLLEDTEPIDTITIMANREVCVQPEKGRKPAIIVPDLGAVREKALYD
VSTLPOAVMSSYGFQYSPGQREVELVAMWKKKPNFPAFADTCFSTVYESDIRV
ESIIYOCCLAPBAROAIIRSLERLYIGPLTNSKQMGXNPSRCASTGLTSCNLT
TCYLKASACRAKLODCTMLVCGDVLVIGESAGTCDASLRVETAMRYSPG
DPRREYDLELTTCSSSNVSVNADSGRQVYVLTDPDPTTLARAMWERTARHPVSM
GNIMVAPTLMARMLTHMTPSITLAAQBLEALDCQITGACTSIQEDLPDITDRLH
GLSAFSLHSYSPGEINRVAACTKRLGVPLRVWRHARSVRKXLSLEGRALANCKRL
FNMAVRIKXKLPPIPAASQLDLSWFWVAGYSGDIIYHSLSRARPRWFMLCILLLSVGV
GIYLLPNR"

ORIGIN

Query Match 87.0%; Score 900.6; DB 14; Length 9359;
Best Local Similarity 91.9%; Pred. No. 1.4e-197;
Matches 951; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 1 ACAGAGTGGATGGGATGGCTGCAAGGTAGCGCTCGGCGTGAACACTCTCTACGG 60
DB 6696 ACCGAGGTGGATGGGATGGCTGCAAGGTAGCGCTCGGCGTGAACACTCTCTACGG 6755

QY 61 GAGAGGTACATTCAGAGTGGGCTCAACCAATACCTGTTGGGTGCGACAGTCCATGC 120
DB 6756 GAGAGGTACATTCAGAGTGGGCTCAACCAATACCTGTTGGGTGCGACAGTCCATGC 6815
QY 121 GAGCCGGAACCGGATGTAGAGTGTCTCATCTTCATGCTCACCGACCCCTCCACATCA 180
DB 6816 GAGCCGGAACCGGATGTAGAGTGTCTCATCTTCATGCTCACCGACCCCTCCATATCA 6875
QY 181 GCAGAGGAGGCGTAAGCGGAGGCTGGCGAGGGGGTCTCCGCCCTCTTGGGACGCTTCA 240
DB 6876 GCAGAGGAGGCGTAAGCGGAGGCTGGCGAGGGGGTCTCCGCCCTCTTGGGACGCTTCA 6935
QY 241 GCTAGCCAGTTGTCTGAGCCCTCTTCTGAGAGCGACATACATTAACCAAAATGACTTCCA 300
DB 6936 GCTAGCCAGTTGTCTGAGCCCTCTTCTGAGAGCGACATACATTAACCGATGACTTCCA 6995
QY 301 GACGTGACTTCATGAGGCGCACTCTCTGTGGCGGCAATGAGATGGCGGGGACATTAAC 360
DB 6996 GATGCTGACTTCATGAGGCGCACTCTCTGTGGCGGCAATGAGATGGCGGGGACATTAAC 7055
QY 361 CGCGTGGAGTGAAGAAAGAGGTAGTAATCTCGGACTCTTTCGACCCGCTCGAGCGGAG 420
DB 7056 CGCGTGGAGTGAAGAAAGAGGTAGTAATCTCGGACTCTTTCGAGCCGCTTCGAGCGGAG 7115
QY 421 GAGGATGAGCGGGAAGTGTCCGTCCGCGCGAGATCTCTGCGGAAATCCAGAAATTTCCA 480
DB 7116 GAGGATGAGGGAAGTGTCCGTCCGCGCGAGATCTCTGCGGAAATCCAGAAATTTCC 7175
QY 481 CGAGGATGCGCCGCAATGGGCAAGCCCGGATTAACAACCTCGCTGCTGAGATCCTGGAAG 540
DB 7176 CGAGGATGCGCCGCAATGGGCAAGCCCGGATTAACAACCTCGCTGAGATCCTGGAAG 7235
QY 541 GCCCGGACTGAGTCCCTCCAGTGGTATAGAGGTGCCCACTGACCACTTAAGAAGCCCT 600
DB 7236 GACCCGACTGAGTCCCTCCAGTGGTATAGAGGTGCCCACTGACCACTTAAGAAGCCCT 7295
QY 601 CCTATACCACTTCAACGAGAAAGAGACAGTTGTCTGACAGATCCACGATCTTCT 660
DB 7296 CCAATACCACTTCAACGAGAAAGAGACAGTTGTCTGACAGATCCACGATCTTCT 7355
QY 661 GCCCTGGCGAGCTTGGCAAAAGCCTTTGTGTAGCTCCGAGCCGTGGCGTGAAGC 720
DB 7356 GCTTGTGTAGCTTGGCAAAAGCCTTGGAGCTCCGAGCTCCGAGCTCCGAGCTGAAGC 7415
QY 721 GGCACGGAACGCGCCCTCCGACCAATCTCCGACGAGCGGAGAGAGATCTGAGCTT 780
DB 7416 GGCACGGAACGCGCCCTCCGACCAATCTCCGACGAGCGGAGAGATCTGAGCTT 7475
QY 781 GAGTGTATTCTCCATGCCCCCTTGAAGGGAGAGCCGAGGAGACCCGATCTGAGCGAC 840
DB 7476 GAGTGTATTCTCCATGCCCCCTTGAAGGGAGAGCCGAGGAGATCTGAGCGAC 7535
QY 841 GGGTCTTGGTCTACCGTGAAGAGAGCGGTGAGAGACGTGCTGTCTCGATGCTC 900
DB 7536 GGGTCTTGGTCTACCGTGAAGAGAGCGGTGAGAGACGTGCTGTCTCGATGCTC 7595
QY 901 TACACATGACAGGCGCTCTGATCAAGCATGCGCTGGAGAGAAAGCAAGCTGCCATC 960
DB 7596 TACACATGACAGGCGCTCTGATCAAGCATGCTGGAGAGAAAGCAAGCTGCCATC 7655
QY 961 AACGGTTGAGCAACTTTTGTCTGCTGACCAACAATGATCTACGCTACCAATCCGCG 1020
DB 7656 AACGGTTGAGCAACTTTTGTCTGCTGCTGACCAACAATGATCTACGCTGCTG 7715
QY 1021 AGCGCAAGCCAGCGG 1035
DB 7716 AGCGCAAGCCAGCGG 7730

RESULT 12
AF208024 AF208024 9373 bp RNA linear VRL 30-MAR-2000
DEFINITION Hepatitis C virus strain MD34, complete genome.

ACCESSION AF208024
VERSION AF208024.1 GI:7341102
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepadnavirus.
1 (bases 1 to 9373)
REFERENCE Nagayama, K., Kurosaki, M., Enomoto, N., Miyasaka, Y., Marumo, F. and Sato, C.
TITLE Characteristic of hepatitis C viral genome associated with disease progression
JOURNAL Unpublished
AUTHORS Nagayama, K., Kurosaki, M., Enomoto, N., Izumi, N. and Sato, C.
REFERENCE 2 (bases 1 to 9373)
TITLE Direct Submision
JOURNAL Submitted (24-NOV-1999) Second Department of Internal Medicine, Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-ku, Tokyo 113-8519, Japan
FEATURES
source Location/Qualifiers
1. .9373
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/strain="MD34"
/db_xref="taxon:11103"
/note="from a patient with persistently normal ALT values (ASC-14)"
gene
1. .9373
/gene="MD34"
330. .9356
/gene="MD34"
/codon_start=1
/product="polyprotein"
/protein_id="AA61205.1"
/db_xref="GI:7341103"
/translation="MSTNPKPORKTKRNTNRRODVKPRGGQIVGVYLLPRGPRLLGYRATKTSERSQPRGRPOIPKARQPBGRMAQPGVPMPLVNEGMGAMGLISRGSRNMGPTPRRRSRRLGKVIDLTGTFADIMKTIPLVNGAPLGAARALAHGRVIEDGNVATGNLPGCSFSIFLLALISCLTTPASAEVRNVSQVHTVNDSSIVTEADMTMSPGCVPCVAREGNSRCWALPTPLAARNGSVPTTIRRVLDLVNSAAACSMYVVDLCGSVPLVLSQLEFSPRRHETVDDCSIPGHLSGRMAAMMMMSPTAALVSGLLRPOAVMDVAGHWGVLGALAYSGVGNWAKYLIMLLFAVDDMLGSSQGLTTRRFLSPSGPSONIOLVTEGMMHNRITLNCNDSHOTGTLALFAHKNTNCSGLTERLNSCHPIDKRDQSGPLTYVEHSGSOKEPCRYARPCGIVYASKYVTEFTPEVVVGTIDREAPLYTNGENETDVLNLNTRPQGMNFGCTWMSSTGYKCGAPCNIGVAGNDTLCTPDCFRKHPEATYTKGSGPMLTPRCVNDPYRLMYPCTVNTVFKVMYGVGVNHLISAACNWTGERCNLEDRSELSPLLTTEQIILPSTFTLPLASTGLIHLHONIYDVOYLGVGSAVVSFYIMKEYVLLFLLLADARICAMMMMLIIPQEARPLENIVLNAASVAGAHGILSFVPCAAWYIKGKLVGAVALGVAMPILLLLLALPPRAYAMDREMAASCAGAVFVGLVLTSPYVPLARLIMSQPTTTRAHIQVWPLNVRGGRDAILLTCVAHPELITDKLLAILGLAMVLAQSIIVKVPFVAQGLIRACMLARKVAGHYVMAFMKLAALITGVYDHLTPLODAHAAASLDDAAVAPVVSDDMETKIIITWADTAACGDIILGVPVAREKEILIGPABSLBEOGRMLIATAYSCOTRLAGCIITSLGRDNQVEGEVOVSTATOSFLATCVNVEGMEVTHGASGKTLACGKPIITQMTYNTNDDLVGMQAPGARSLTPTCTGSSDLVYVTRADVIYPRRGDSRGSLSPPRVSYLKSSGGPILCPSGHANGITPRAVCTGVAANDPVYBESHTMRSPVFTDNSSPPAVPQTFQVAHLHAPFGSGSKTRPAVAAQGTIVLVNPSVAATIGFAYMSKVAHGVDPNIRTGARTVTGAPITYSTYKFLADGCGSAGADIIMCECHSDTSTIIGITVLDQAEAGARLVVLAATATPGSVTPHPNIEVALNTGEIIPYGAIPLETIKGRHLLIFCHSKKCDLALGSLGNAVAYVGLDVSIPASGNVAVVATDAMTGVTGDPDSVLDICNTCVTQVDFSLDPTLETITTVPGADVSRSGRGRGRTYRTPYRBERPSGMDSSVLECEYAGCAWIELTPAESVRLARVILNTPGAPCYCDHLEPMEVFTGLTHIDAHFLSOTKQAGDNPPVAVOATVCARQAPSPMDQWKLIRKPTLHGPPLPLRLGAVONEVTLTHPTIKYIMACMSADLEVNTSWLVGMLALAAVCLTGSVVIVIGRIISLSPAVIPDREVLKEPEDEMECHASLIEGQMLAOPKOKALGLQTAOTAEVAAPVSKRALEAFKAMMNPISGLIYLAGISTLGPAPIASMAFTSISPLTOSTLLENLIGLAPASALAVGAGVGAIVCAITLKVYDILAGYAGVAGALVAPKVMGEMPTEDLVNLLPALSBAVLVYCAAILKRVVGBEAGVQWNRLLIAPASRGNHVSPTHVPSDAAARVQIISLITVQLRLHOMINDCSTPGSGWLRVDWMICTVLDFKTMLOSKLPRLPVGPSPSCGYSFAPWMDGIMQTCPCGAQITGVNKSRIIGPKTCSNTHTGTFINAYTGPCTSPAPNYSRALMRVAABEYEVTRVGDPHYDKPCPDVPAEFTEVDVGLHAYAPACKPLLRBEVTFQVGLNQLVNSQLPCEBEPDVAVLTMLTDPHSITBAAKRRLAR

ORIGIN

Query Match 87.0%; Score 900.6; DB 14; Length 9373;
Best Local Similarity 91.9%; Pred. No. 1.4e-197;
Matches 951; Conservative 84; Indels 0; Gaps 0;

QY 1 ACAGAAAGTGAGTGGGTGCGGCTGCACAGATACGCTCCGCGTCAAACTCTCTTAACG 60
Db ACAGAAAGTGAGTGGGTGCGGCTGCACAGATACGCTCCGCGTCAAACTCTCTTAACG 6749
QY 61 GAGGAGTCAATTCAGAGTCCGGCTCAACCAATACCTGGTGGGTGCGAGCTCCCATGC 120
Db GAGGAGTCAATTCAGAGTCCGGCTCAACCAATACCTGGTGGGTGCGAGCTCCCATGC 6809
QY 121 GAGCCCGAACCAGATATAGCAAGTCTCATCTTCATGCTACGAGACCCCTCCCATCACA 180
Db GAGCCCGAACCAGATATAGCAAGTCTCATCTTCATGCTACGAGACCCCTCCCATCACA 6869
QY 181 GCAGAGACGCTAAGCGCAGGCTGCGCAGGAGGCTCTCCCTCTTGGCCAGCTCTTCA 240
Db GCAGAGACGCTAAGCGCTGCGCAGGAGGCTCTCCCTCTTGGCCAGCTCTTCA 6929
QY 241 GCTAGCAATGTGTCTGCGCCCTTCTCTGAGAGGCACTAATTAACCAATGACTTCCCA 300
Db GCTAGCAATGTGTCTGCGCCCTTCTCTGAGAGGCACTAATTAACCAATGACTTCCCA 6989
QY 301 GACGCTGACCTATCGAGGCAACCTCTGTGCGGCGATGAGATGGGCGGAGCATTTACC 360
Db GACGCTGACCTATCGAGGCAACCTCTGTGCGGCGATGAGATGGGCGGAGCATTTACC 7049
QY 361 CGGCTGAGTCAAGAAACAAGTAGTAATCTTGAATCTTTGACCCGCTCCGAGCGAG 420
Db CGGCTGAGTCAAGAAACAAGTAGTAATCTTGAATCTTTGACCCGCTCCGAGCGAG 7109
QY 421 GAGGATGACGGGAGAGTGTCCGTCGCGGAGATCTCTCGGAAATCCAGAAATCCCA 480
Db GAGGATGACGGGAGAGTGTCCGTCGCGGAGATCTCTCGGAAATCCAGAAATCCCA 7169
QY 481 CCAGCATGCCGCGATGAGGCAAGCCGATTCCTGTGCGGAGATCTCTCGGAAATCCAGAAATCCCA 540
Db CCAGCATGCCGCGATGAGGCAAGCCGATTCCTGTGCGGAGATCTCTCGGAAATCCAGAAATCCCA 7229
QY 541 GCCCGGAGTACGTCCTCCAGTGTGATAGGAGTCCCACTGCGCATCTTAAGACCCCT 600
Db GCCCGGAGTACGTCCTCCAGTGTGATAGGAGTCCCACTGCGCATCTTAAGACCCCT 7289
QY 601 CCTATACCACTTCACAGAGAAAGAGACAGTTGTTCTCAAGAAATCCACCGTGTCTTCT 660
Db CCTATACCACTTCACAGAGAAAGAGACAGTTGTTCTCAAGAAATCCACCGTGTCTTCT 7349
QY 661 GCCCTGCGGAGCTTGCACAAAGCTTTGGAGTCCGAGACCGTCCGAGCTTCAGAC 720
Db GCCCTGCGGAGCTTGCACAAAGCTTTGGAGTCCGAGACCGTCCGAGCTTCAGAC 7409
QY 721 GGCACGGCAACGCGCCCTCTTGAACCAATCTTCCAGACGAGCGGAGACAGATCTGAGCTT 780
Db GGCACGGCAACGCGCCCTCTTGAACCAATCTTCCAGACGAGCGGAGACAGATCTGAGCTT 7469
QY 7410 GGCACGGCAACGCGCCCTCTTGAACCAATCTTCCAGACGAGCGGAGACAGATCTGAGCTT 7469

Qy	781	GAGTCGATTCCTCCATGCGCCCTTGAAGGGGAGCCCGGAGACCCCGATCTCAGCGAC	840
Db	7470	GAGTCGATTCCTCCATGCGCCCTTGAAGGGGAGCCCGGAGATCCGATCTCAGCGAC	7529
Qy	841	GGGTCTTGGTCTACCGTAGTAGGAGCGCGGTGAGAGAGTGTCTGCTGCTGATGTC	900
Db	7530	GGGTCTTGGTCTACCGTAGTAGGAGCGCGGTGAGAGAGTGTCTGCTGCTGATGTC	7589
Qy	901	TACACATGAGACAGCGCTCTGATCAGCCCATGCGCGGAGAGAAAGAGTGGCCATC	960
Db	7590	TACACATGAGACAGCGCGCTCTGATCAGCCCATGCGCGGAGAGAGAGTGGCCATC	7649
Qy	961	AACGCGTTGAGCAACTTGTGCTGCTGCTCAACCAACATGATCTACCGCTACCAATCCCG	1020
Db	7650	AATGCGTTGAGCAACTTGTGCTGCTGCTCAACCAACATGATGATGCTACCAATCCCG	7709
Qy	1021	AGCGCAGCGACGCG 1035	
Db	7710	AGCGCAGCGACGCG 7724	
RESULT 13			
LOCUS	AF207753	9374 bp	RNA linear VRL 27-Apr-2000
DEFINITION	Hepatitis C virus strain MD12 complete genome.		
ACCESSION	AF207753		
VERSION	AF207753.1	GI:7650223	
KEYWORDS			
SOURCE			
ORGANISM	Hepatitis C virus		
REFERENCE	1 (bases 1 to 9374)		
AUTHORS	Nagayama, K., Kurotaki, M., Enomoto, N., Miyasaka, Y., Marumo, F. and Sato, C.		
TITLE	Characteristics of hepatitis C viral genome associated with disease progression		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 9374)		
AUTHORS	Nagayama, K., Kurotaki, M., Enomoto, N., Miyasaka, Y., Izumi, N. and Sato, C.		
TITLE	Direct Submission		
JOURNAL	Submitted (23-NOV-1999) Second Department of Internal Medicine, Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-ku, Tokyo 113-8519, Japan		
FEATURES	Location/Qualifiers		
source	1..9374		
	/organism="Hepatitis C virus"		
	/mol_type="genomic RNA"		
	/db_xref="taxon:11103"		
	/note="Isolated from a patient with persistently-normal ALT values (ASC-2)"		
	330..9362		
	/product="polyprotein"		
	/protein_id="AAF65943.1"		
	/db_xref="GI:7650224"		
	/translation="MSTNPKPQRTKRNTRPQVYKPPGGQIYGVYLLPRRPRRLGVRATKTSERQPRGRQPIKVRQPBKRTWAQPGYEMPLPGNGKMWAGLLSPRGSRPSWPTDPRKRSRLKQVLDITLTCGADLNGYIPLVGAIPGCGARVLAHQVFLVDGMYATGNLPGCSFSIPLIALSLCTTPASAEVNVAGMYHVDLNDGNSIYVLAADMTMAHAGVCPVARENSSRCWALTPTLAARASVPTTIRHVDLVGAAGFCGMYVGDLCGVFLISLFTPSPRHETVODNCISYPGHVGHMAMMMMMNSPTALVY SOLAIRPQAVVDVAGAHGVTLAGLAAYISMAGNNAKVLWLLPFGVDGQTRISGCTOCHTISGLVLSRSSGSQNLQLININGSHWINTALNCDSLNTGFLALFTHKNASGCPARMASCRSIDTFAQMGPIYTAQPSISDQRCMYAPQCGIIVASQGVKVCYFPSPVVVGTDRSGVPTYSWGENETDVLINNTPRPQGMNPGCTMANSTGTCKCGPCPNIGVGNNTLTCPTDCFRKHPEATYTKCGSGPMLPRGLVDVYRLMHPCTKVF SIFKRYMVGVGEHRLNAACNWTBGRGLEDGRSELSPLISTTEMOU.PCSPTTL PALSTGLIHHQNVVDVQIYIGISGAVVYSFYVKNDYVVLPLILADAVACACMMMLL IIAQEAALBNLVVLAASVAGAHGFLSVFFCAAWYIKGRLVPGAALYALGVWFLLL IIALUPPRAVYAMDREMAASCGGAVLVGMILITLSPYKVFILARLLIMLQYILITRAEHL		

Qy	1	ACAGAGTGAATGGGGTGGCTGACAGAGTACGCTCGGGGTGCAACCTCTCCTACGG	60
Db	6636	ACTGAGGTGAATGAGTGGCTGACAGAGTACGCTCGGGGTGCAACCTCTCCTACGG	6755
Qy	61	GAGGAGTCAATTCAGAGTGGGCTCAACCAATACCTGTTGGGTGCGACCTCCATGC	120
Db	6756	GAGGAGTCAATTCAGAGTGGGCTCAACCAATACCTGTTGGGTGCGACCTCCATGC	6815
Qy	121	GAGCCCGAACCAGATGTAAGAGTGTCACTTCATGCTCAACCGACCTCCCAATCA	180
Db	6816	GAGCCCGAACCAGATGTAAGAGTGTCACTTCATGCTCAACCGACCTCCCAATCA	6875
Qy	181	GAGAGAGCGGTAAAGCGCAGGCTGGGCGAGGGGTCCTCCCGCTCTGGGCGAGCTTCA	240
Db	6876	GAGAGAGCGGTAAAGCGGAGGTTGGCGAGGGGTCCTCCCGCTCTGGGCGAGCTTCA	6935
Qy	241	GCTAGCCAGTTGTCTGGCCCTCTCTCGAAGCGCATATCAATTAACCAATGACTTCCA	300
Db	6936	GCTAGCCAGTTGTCTGGCCCTCTCTCGAAGCGCATATCAATTAACCAATGACTTCCA	6995
Qy	301	GACGCTGACCTCATGAGGCGCAACCTCTGTGGCGGCGATGAGTGGCGGGAACATTAC	360
Db	6996	GACGCTGACCTCATGAGGCGCAACCTCTGTGGCGGCGATGAGTGGCGGGAACATTAC	7055
Qy	361	CGCGGAGTGAAGAGCAAGAGTGTATCTCGAAGCTTTTCGACCGCGCTCGAGCGGAG	420
Db	7056	CGCGGAGTGAAGAGCAAGAGTGTATCTCGAAGCTTTTCGATCGGCTCGGCGGAG	7115
Qy	421	GAGATGAGCGGGAAGTGTCCGTCCCGCGGAGATCTCGCGGAATCCAGAAATTTCCA	480

ORIGIN

Query Match

Best Local Similarity 86.1%; Score 891; DB 14; Length 9374;

Matches 945; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

1

6636

61

6756

121

6816

181

6876

241

6936

301

6996

361

7056

421

LOWVPPLNVBGRDAILFLTCATHEPILFDITKLALLGLMLVLAQGITRVFVR
AAGLIRACMLRVKTYGVYVOMAFKALALGTYVYHDLTPLRGAFHLDLVAVE
PVPSDMETKTIITMGADPAACGDIILGIPARRGRRLILDPALPSBOGRILAPIT
AYSQOTRLILCIIITSLTGPRNOVBGVQVYVSTATOTFLATCIVGCVTPRHAGSK
TLAEGKPIITMYTNVDDLVGQWAPBARSMPTCTGSSDLVAVTRADVI PVRRG
DGRSLSPRPVYSILKSGSGGLCPGSHVVGII FRAACITGVAKAVADPVEVSEMT
MRSFVFTDSSPPVAPQTFQVAAHLAPGSSKSTKVPVAAQCYKVLVNVSVATL
FGAYMSKAYSGDNRIRTVARTITGAPITTYTGKFLADGCGGAGADIIICBCHS
TDSTIIIGITVLVAGVAGARLVLAIVAPRGSVTVPVPIREVALSNGETIPYK
AIPETITGSHLIFCHSKKCDLAKLAKLGINAVAYRGLDVSYPISGDIYVVA
TDALMTGTFDPSVIDNTCVTPVPSLDPITDTITVTPQVADSVTORRGRRG
RGIYRVTPEBRSRSGMDSVLCEBQDAGAWELPETSRLRALNTPGLPVC
DLHFEWEGVFTGLRIDAHFLSOTKQAGDNPPYVAYOATCAABAPPSMDMKC
LTRKPLHAGPTPLTKYIMTCMSADLVEVTSVTLVGGVLA
AALVCLTGSVIVGRITISGRALIPDEVLVIREPMECSOLPYIBOGMOLAE
QPKOKALDLTATKQABAAAPVYBSKKRALEARGKMMNFISGITYLALSTLPCN
PALASLMAFLTAGSVTSPLTQSTLLFNILGAVNAQAPASASAFVGGIAGVAGSI
GLGKVLVDILAGYAGVAGALVAFKMSGEMSTEDLVNLPLALISPGALVGVCA
ILRHVGEAGAVOMNRLLIAPASRGHVSPTHVYESDAARVYQILSNLTQLLR
RLHOMINDEGTPCSGSLRDVMDMICTVLSDPFKMLQSKLPLRPGVPLSCRGYK
GWRGDGIMOTTCGCAQITGHWNGSWRIYGPRTCSATWHTGTPRINAYTTPGCTSP
APNYSBALMRVAEEVYVTVYCDPHYTGCTDNVCKPCVPAPBEPFTVDGRRLR
YAPACKPLREEVTFQVGLNQVLVGSQULPCPEEDVAVLITSLTDPSTITAEAKRL
ARGSPSLASSASQULSAPLAKCTTHDPPADLLEANLMLQEMGNITRVESN
KVLILDSFDPARAEDEIREISVPAEILIRKPKFPALPTWAPDPYNPLBSMDPDY
VPPVHGCPLPPTGKPLPPRRKRTVLTTESTYSALAEATRTFGSGSADSGT
ATAPDOSTDGDGSDVESYSMPELEGEEDDLSGMSSTVSEERAGEVCSWS
YTTGALITTPCAABESKLPINALNSILRHNNMYATTSRACOROKYTTDRLOVLD
DHYRDLKEMKAKASTVAKLISVEACKLPPHSASRFGYAGADVNLSKAVNHI
RSVKDLLEDDETETINTVMANKNEVFCQPEKGRKAPRLVFPDLGRVCEKALYD
VESITLPOAVMGSYSFOYSQGVREFLVNTWKSKKCPMGFSYDRCEPSTIENDIRV
EESIYOCDDLAAPBARVKSRLTERLYVGGPLTNSKGQNGYRCARCSVLTCGNTL
TCYLKASAGCAALLOCTMANGDILVIVICESAGTODASLIVFEEMARVYAPG
DPPEPVDLEITTCSSSVSYAHNDASGRVYITRDPPTPLARAMEATRRTPNWSL
GNITMIFATLMARMLMTHFPSILADQSKYALDCQYGAQCSIEPLDQITERHL
GLSAPLSHSYSPGEINRVASCLRLKGLVGPPLVGMHRRASVLAISOGGRATQSKYL
FNMAVRTLKLTLPAPASRLDLSGMFVAGYSQIYHLSRABRBRMFWLCLLLSVGV
GYLLPNR"


```

Db      7116 GAGGATGAGAGGAAATATCCGTTCCGGCGGAGATCTCTGGGAAACCCAGGAAGTTCCCC 7175
Qy      481  CCAGCATGCCCGCATGGGACGCGCGGATTTACACCTTCGCTGTGAGTCTTGGAG 540
Db      7176 CCAGCGTTCCTTATGGGACGCGCGGATTTACACCTTCGCTGTGAGTCTTGGAG 7235
Qy      541  GCCCGGATACGTCCTCCAGTGTACATGGGTGCCCATGCGCCTACTAGACCCCT 600
Db      7236 GACCCGAGTACGTCCTCCGCTGGTGCAGCGGTGCCATTTGACCTTACCAAGGAGGAGCT 7295
Qy      601  CCTATACCACTCCAGGAGAAAGAGACAGTGTCTGTGACAGATTCACCGTCTTCT 660
Db      7296 CCATTACCACTCCAGGAGAAAGAGACAGTGTCTGTGACAGATTCACCGTCTTCT 7355
Qy      661  GCCCTGGCGGAGGCTTCCCAAAAGCTTTGTAGTCCGGAACGTCGCGCGTGAACAG 720
Db      7356 GCCCTGGCGGAGGCTTCTACTAGACCTTTGGCAGCTCCGGGTCGTCCGCGCGACAGT 7415
Qy      721  GGACGCGGCAACGCGCTCTCTGACCAATCTCCGAGACGCGCGGAGACGATCTGACGTT 780
Db      7416 GGACGCGGCAACGCGCTCTCTGACCAATCTCCGAGACGCGCGGAGACGATCTGACGTT 7475
Qy      781  GAGTGTATTCCTCCATGCGCGCTTGAAGGGGAGCGCGGAGACCGCGATCTCAGGAGC 840
Db      7476 GAGTGTATTCCTCTCCATGCGCGCTTGAAGGGGAGCGCGGAGACCGCGATCTCAGGAGT 7535
Qy      841  GGGCTTGTGCTACCGTGAAGAGAGCGCGGTGAGAGACGTCGTCTGCTGCTGATGTC 900
Db      7536 GGGCTTGTGCTACCGTGAAGAGAGCGGTGAGAGACGTCGTCTGCTGCTGATGTC 7595
Qy      901  TACACATGAGACAGGCGCTCTGATACGCCATGCGCTGCGGAGAAAGCAAGTCCCATC 960
Db      7596 TACACATGAGACAGGCGCTCTGATACGCCATGCGCTGCGGAGAAAGCAAGTCCCATC 7655
Qy      961  AACGCGTTAGCAACTCTTTGCTGCGTCAACCAACATGATTAACCATCCCGC 1020
Db      7656 AACGCGTTAGCAACTCTTTGCTGCGTCAACCAACATGATTAACCATCCCGC 7715
Qy      1021 AGCGCAAGCGACGGC 1035
Db      7716 AGCGCAAGCGACGGC 7730

RESULT 14
AF165059 9379 bp RNA linear VRL 04-NOV-1999
LOCUS      Hepatitis C virus strain MD8-1 complete genome.
DEFINITION
ACCESSION AF165059
VERSION    AF165059.1 GI:5918956
KEYWORDS
SOURCE     Hepatitis C virus
ORGANISM   Hepatitis C virus
Virusess: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepadnavirinae.
REFERENCE 1 (bases 1 to 9379)
AUTHORS   Nagayama,K., Kurotsaki,M., Enomoto,N., Maekawa,S.Y., Miyasaka,Y.,
Tazawa,J., Izumi,N., Marumo,F. and Sato,C.
TITLE     Time-related changes in full-length hepatitis C virus sequences and
hepatitis activity
JOURNAL   Virology 263 (1), 244-253 (1999)
MEDLINE   20013325
PUBMED   10544098
REFERENCE 2 (bases 1 to 9379)
AUTHORS   Nagayama,K., Kurotsaki,M., Enomoto,N., Maekawa,S., Miyasaka,Y.,
Sakamoto,N., Fukuma,T., Tazawa,J., Izumi,N., Marumo,F. and Sato,C.
TITLE     Direct Substitution
JOURNAL   Submitted (06-JUL-1999) Second Department of Internal Medicine,
Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-ku,
Tokyo 113-8519, Japan
FEATURES
source    1..9379
/organism="Hepatitis C virus"

```

CDS

```

/mol_type="genomic RNA"
/strain="MD8-1"
/db_xref="taxon:11103"
/notes="from a patient with persistently high serum ALT
values; patient 8, point 1
genotype: 1b"
330..9362
/codon_start=1
/product="polyprotein"
/protein_id="AAD56194.1"
/db_xref="GI:5918957"
/translation="MSTNPKPQKTKRNTNRAPQVFKPGGQIVGVLLPRGRPL
GVATRKTSERSPQPRGROPKIPKARPERSTAAQGYPMPLVNGEMGMWLSRSG
SRPSMGNDPRRPSRNLCVITDTTCGPRADLMGYPVAGAPGCAARALAHVRVLE
SVNVAITGNLPGCSFSIFLALISCLTIPASAEVKNVSGVHTVNDSCNSIYEAAD
IIMHPGMPQVREDNISRCVVALPTLAAANI SVPTTIRHVLVLGTAAFCSAMV
VVDLGSFVLVLFPSPRRHETVODCSIPGHVSGRMMMDMMSPALAV
SOLLRPOAVDMVAGAHVGLAGLAVSVGNMAKVLIMLLFAGVDGTTVTGGTO
ASAHILTSFLTVGPSOKIOLVNTNGSMHINRTALNDSILQTFALALFYHRFNAS
GCPERMASSRPIDTPOGHWPTTYAPENNPDRPYCMHAPRPGCIERASKCPYVC
FTPSFVVGTGTDREGVPTNMGENSEDTVLNNTRPQGNWFGCTMNSTGTCTCGG
PECNIGAGANNLLICPTDCFRHNPATYTKCGSGWMLPRCLVDYPRILMHPCTVNF
TIFKRMVYGVGEHRLSAAQMWTRGERCDLEDRDSELSPLISTTEQIILPCSFPTL
PALSTGLHLHRNIIDVOYLXGVGAAVSFAIRMEYVLLFLPLADARCAQIMMLL
IDARBALENIIVLNAASVAGAHGISLFLFPCCAAMYIOGRLVPGAAVAVGVMPPLL
LILALPRRAYAMDREMASSCGGAVVGLAFILSLPYRYFLARLIMWLOYPTTRAEH
LDVWPPLNVRGGRDAILILMCVHRELFDITKLILFGLMVLQAGITVAFVFR
AAGLIRACMLVRKAVAGHVQMAFVKALALNTGYVLDLAPLQMAHSAQRLAAVAV
PVPFSDMETKILITWGDADPACGDIILGLPVSARRREILIGPADSLGCGMPLAIT
AVSOOTRMLGCIITSLNDRGNQVEGVEQVSTVOSPLACINCGCTVPHGASK
TLAGKRPITOMYTNVTDOLVGMQAPGASMTPTCCSSDLVLTVRHADVI PVRRG
DGRGSLSPRPVSYLKSSGCGPLLPDSGHAVIDPAAVCTRRVAAVDPVPRESMETT
MRSVPFTNDSPPAVPQTEOVHLHAPFSGSTVPAPVAAQKRVLVNLSVAVETL
GGGAWSKAYGTDPMIRGITRTITGAPITVSTYKFLPADGCGGAGADIIICDECHS
TDSITLIGIGTLDQAEAGARLVVILATIPGSGTVPPHPIEVALSTGIBIPFGH
APIPIVIGKGRHLIFCHSKKCDLEAKLSALANAVAYRGLDVSLPSTGADVVA
TDALMTGYGDEDSVINDCTVOTVDVSDLSLPTFTIETTVPQDAVSRGRRGRLR
RGCIPTPTPRGSRPGMPSVLCRCYDAGCAWYLTAPETSVRLAYRTMGLPQPC
DHLEFMESEFTGLTHIDAFILSQTKQAGNFFVLAQVQCARQAPPSMDQMKK
LIRLRKPTLHGPPLLYRLGAVQNEVTLTHPIKFKIMACNSADLEVSTVVGVLAA
ALAAVCLTTSQVIVIRIILSGKPAVTPREVLVREPDMECEASHLPITBQMOLAE
OFKOKALGLQATKOAEAAAPVSEKRALETPAKKHMNFISGIVLAGISTLPGN
PALAISMLFTASITSLPTTOTTLMENIIGGVAALAPASASAPAGVADIAAAGVSI
GLGKLVLDLAIAGYAGVAGALVAFVMSGEMSTEDLILPALISPGVLYGCAAI
ILRRVGRBGAIVQMMNRLIAPASRGNHSPHYIPESDAAKAVQIUSLITLQILK
RLHQWINDCSTPSCGSMRLQDKWIMICVLSPFKWLQSKIMPLRGVFLSCQRQYK
GVMRGDGMQITPCGAQITGHVKNKSMRYGPKTCSNMHGTFPINATVTPCTPSY
AVNYSRALMRVAAEYEVTRVGDHYVGTMTDVKCCQVAPAEFTFVGVRLHR
VAPACKPLIREVTVQVGLNOVLYGSOILPCBEPDVAVITSLMIDPSHITATKARL
ARGSPSLASSASQLSAPSKATCTTHHSDADLILANLIMROBMKGNITRVESN
KVVILDSFDPLAEEDERBVAAILRRSRKFPALPVMAIPDNPPLBSKMDPDY
VPPVHVGCPLEPTKAPPIPPRRKRTVLTSTVSALAEKATKTFSGSGSAAVSDGT
ATAAPDOASDDDESDVESYSMPLKEEPPDLSGDSMTVEEASEDVCCSMS
YTMGALITPCDAESKLPINPLNSLIRHNMVATTSRSASLQKQVTPRLQVLD
DHYRVULKEMKAKATVYAKLISIEACKTPRPARBKGFGYADVSISSKAYNI
RSWMDLLEDTETPIDTITMAKSEVFCQPEKGGKRPALIVPDLGRVRCMALYD
VSTLIPQAVMSGSGYFQYSPKQVPEPLVNTWKSCKPMPGFSYDTCPCSTVYESIRT
TESIIYOCCLDAPEARKAKISLTERLYIGGLPLNSNGOAGYRCARSAEGLTYSNCLT
TCYLKASACRAKAKODCTMLVNGDVLVIGCSAGTQEDAAALRYATEMTGSAAPG
DPDPOPEDELITSCSSNVSVAHDSAGRYVYLTDPPTPLARAMEYRBPVNSWL
GNIMVAPTLMAKMLTMHFPESIIIAQOLEKALDCQIYGCYSYIETDIPOLIRRH
GLSARSLHYSFPEINRYASCDKRLGVPLRVWRHRASVYRAKLISQGRATTCGKYL
FNMAVRTKLKLTPIPAASRLDLSGWFVAGYSGDITVHLSLRARPRPMWMLLISVGV
GIYLLPNR"

```

ORIGIN

```

Query Match      86.1%; Score 891; DB 14; Length 9379;
Best Local Similarity 91.3%; Pred. No. 2.4e-195;
Matches 945; Conservative 0; Mismatches 90; Indels 0; Gaps 0;
Qy      1  ACAGAGTGTAGTGGGTGGCGCTGCACAGGTACGCTCCGCGCTGCACAACTCTCTTACCG 60
Db      6696 ACAGAGTGTAGTGGGTGGCGCTGCACAGGTACGCTCCGCGCTGCACAACTCTCTTACCG 6755

```

61 GAGAGGTGCAATTCCAGGTGCGGGCTCAACCAATACCTGTTGGTGCAGCTCCCATGC 120
6756 GAGGAGGTGCAATTCAGGTGCGGGCTCAACCAATACCTGTTGGTGCAGCTCCCATGC 6815
121 GAGCCGGAACCGGATGTAGCAAGTGTCTACTTCCATGCTCAACGACCCCTCCCATGACA 180
6816 GAGCCGGAACCGGATGTAGCAAGTGTCTACTTCCATGCTCAACGACCCCTCCCATGACA 6875
181 GAGAGAGCGGTGAAGCGAGGCGTGGCGAGGGGGTCTCCCGCTCTGGGCGAGCTCTTCA 240
6876 GCGAGAGCGGTGAAGCGAGGCGTGGCGAGGGGGTCTCCCGCTCTGGGCGAGCTCTTCA 6935
241 GCTAGCACTTGTCTGGGCTCTTCTTGAAGGCGACATACATTAACCAATGCTTCCCA 300
6936 GCTAGCACTTGTCTGGGCTCTTCTTGAAGGCGACATACATTAACCAATGCTTCCCA 6995
301 GACGCTGACCTCATCGAGGCGCAACCTCTGTGGCGGCGATGAGATGGGCGGGAATTAC 360
6996 GACGCTGACCTCATCGAGGCGCAACCTCTGTGGCGGCGATGAGATGGGCGGGAATTAC 7055
361 CGGTGGAGTCAAGAGCAAGGATGATCTGGAATCTTTGACATCCCGCTTGAAGCGAG 420
7056 CGGTGGAGTCAAGAGCAAGGATGATCTGGAATCTTTGACATCCCGCTTGAAGCGAG 7115
421 GAGATGAGCGGGAAGTGTCCGTCCCGGCGAGATCTGCGGAATCCAAAGAAATTTCCA 480
7116 GAGATGAGAGGAGGAGTATCCGTGGCGGGAATCTTGGGAGATCCAGAAAGTTCCCC 7175
481 CCAAGCATCCCGCATGGCGAGCGCCCGGATTAACACCTTCGCTGTGGAGTCTGGAAG 540
7176 CCAAGCATCCCGCATGGCGAGCGCCCGGATTAACACCTTCGCTGTGGAGTCTGGAAG 7235
541 GCCCGGAGTACGTCTCCAGTGTATCGAGTGGTCCCATCGCACTTAAGACCCCT 600
7236 GACCCGAGTACGTCTCCCGGAGTATCGAGTGGTCCCATCGCACTTAAGACCCCT 7295
601 CCTATACCACTTCAACGAGAAAGAGAGATGTTCTGACAGAAATCCACCGTCTTCT 660
7296 CCAATACCACTTCAACGAGAAAGAGAGATGTTCTGACAGAAATCCACCGTCTTCT 7355
661 GCCCTGGCGAGTGTCCCAAAAGCTTTTGGTGAAGTCCGGAACCGTGGCGCTGACAGC 720
7356 GCCCTGGCGAGTGTCCCAAAAGCTTTTGGTGAAGTCCGGAACCGTGGCGCTGACAGC 7415
721 GCGAGGGAACCGCCCTCTGACCAATGCTCCGAGAGCGGCGAGCAAGATCTACGTT 780
7416 GCGAGGGAACCGCCCTCTGACCAATGCTCCGAGAGCGGCGAGCAAGATCTACGTT 7475
781 GAGTGTATTCCTCATGCGCCCTCTTGAAGGCGAGCGGCGAGCAAGATCTACGCGAC 840
7476 GAGTGTATTCCTCATGCGCCCTCTTGAAGGCGAGCGGCGAGCAAGATCTACGCGAC 7535
841 GGGTCTTGGTCTACCTGATGAGAGCGCGGATGAGAGCTGTGCTGCTGATGCTC 900
7536 GGGTCTTGGTCTACCTGATGAGAGCGCGGATGAGAGCTGTGCTGCTGATGCTC 7595
901 TACACATGACAGGCGCTGATGACGCGATGGCTGGAGGAGAAAGCAAGCGCCATC 960
7596 TACACATGACAGGCGCTGATGACGCGATGGCTGGAGGAGAGCAAGGTTGCCATC 7655
961 AACGGCTTGAAGCACTTTTGTGCTGACCAACAATGCTTACGCTTACCAATCCCGC 1020
7656 AACGGCTTGAAGCACTTTTGTGCTGACCAACAATGCTTACGCTTACCAATCCCGT 7715
1021 AGCGCAAGCGAGCGG 1035
7716 AGCGCAAGCGCTGGG 7730

RESULT 15
API65060
LOCUS AF165060 9379 bp RNA linear VRL 04-NOV-1999

DEFINITION Hepatitis C virus strain MD8-2 complete genome.
ACCESSION AF165060
VERSION GI:5918958
KEYWORDS Hepatitis C virus
SOURCE Hepatitis C virus
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
REFERENCE 1 (bases 1 to 9379)
AUTHORS Nagayama,K., Kurosaki,M., Enomoto,N., Maekawa,S.Y., Miyasaka,Y.,
Tazawa,J., Izumi,N., Marumo,F. and Sato,C.
TITLE Time-related changes in full-length hepatitis C virus sequences and
hepatitis activity
JOURNAL Virology 263 (1), 244-253 (1999)
MEDLINE 20013325
PUBMED 10544098
REFERENCES 2 (bases 1 to 9379)
AUTHORS Nagayama,K., Kurosaki,M., Enomoto,N., Maekawa,S., Miyasaka,Y.,
Sakamoto,N., Fukuma,T., Tazawa,J., Izumi,N., Marumo,F. and Sato,C.
TITLE Direct Submission
JOURNAL Submitted (06-JUL-1999) Second Department of Internal Medicine,
Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-ku,
Tokyo 113-8519, Japan
FEATURES
source
1..9379
Location/Qualifiers
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/strain="MD8-2"
/db_xref="taxon:11103"
/note="From a patient with persistently high serum ALT
values; patient 8, point 2
genotype: 1b"
330..9362
/codon_start=1
/product="polyprotein"
/protein_id="AAD56195.1"
/db_xref="gi:5918958"
/translation="MSTNPKPQRTKENTNRNRPDVKFGGGQIVGGVLLPRGRPL
GVRAKTSERSQPRGRPIPKARREGRMAQPGWMPVIGNMGAGMGLSPG
SRPSPGNDPRSRNIGKVIDTLCPADIMGYIPLVGAIPGAALAAAGVAVLD
SVNATGNLPGCSFSTFLALISGLTTPASATYERANSGYHTVNTDSSNSIVYBAD
IIMHTPGCTPCVREGNISRCVVALPTLAAANISVPTTTLRRHDLVGTAFCSANV
VGDLGSEFVLSQFTFSPRHETVQDCNSIYGHSGHGMADMMNNSPTALVY
SOLIRIPOAVVADAGAHMGVLAALAYSGMAKVLIMWLIFAGVGTTHVGGAG
AHSAAHLSMPTTGVRSOKIOLVNTNGSNHIRTALNCNDSLQTFEALAFTHFENS
GCPBRMSCRPIIDPDGKGPITTAENSPDPRCYHVARPGGIPASVGGPPVCS
FTSPVYVGTDRFCVPTTYMGENETVLLNTRPPQGNFECTMANSSTFTTCGS
PPCNIIGAGANTLTCPTDCFRKHPERATYTKCGSPWLPRLCLVDYPRLMYPCVNP
TIFVRYVVGGEVHRLSAACNMTGERGCDLDRDRSELSPLLSTTMOILPGRPTL
PALSTGLIHLEHNIIVDOYLYGVGSAAVSIIRWEYVLLFLILADARVACIMMLIL
IAOAEALLENLVNNAASVAGAGHILSFVFFCAAMYIKGRILVGAAYALYGVNPLL
LIWVPRAYVMDREMASSCGAFVGLAFLTSPYKVFPLARLIIMLYETITPBAH
LQWVPPLANRAGDAIILMCAVHPHELIPDITLLALIREPLMVLQAGITRVYFAR
AHGILRACMLVRKAVGAYVOMAFVKLAALTYGVYVHLARLQHMASGLDLVAAYR
PVSQDMEYKLIITWADTAAGDIIILGIPVARSARGRIILGPAISLGGQGRALAPIT
AVSQOTGGLGCIITSLTGDKNDKEVEVQVSPATOSPLATCINGCMTVFRHAGSK
TLAEPKPIITOMYTNPODLYGMOAPGASMTCTCGSSDLYLVTHADVIYPRRG
DGRSLISPREVSVLKSSGGPIICPGHAGVIGIRAVRCRGVAKAVDPIPVSEMETT
MRSVPPTDNSPPRAVPOTROYAHHAATGSGSKSKVPAAYAAOQYKYLINPSTATL
GFGAYMSKAYGTDNIRGTGRTITGAPITYSTGKFLPDGGCGGAYNDIILICBSH
TDSITLIGITVLDQAEATGARLVVLAATAPGSAVPHPIBEVALSNGEITPYGK
AIPIEVIKGGHILFCHSKKKCDLALAKLSLGLNAAAYYKGLDVSYPISGDAVVA
TDALMTGYTDEPDSVINCNTCVTQTVESLDPPTIETTTVPDPAVSORGRTGG
RGGIYRFTVTPERSGMPDSVLCEDYDAGCAWELPAETSVSLRAYINPGLPVOC
DHLEFMSVFTGLTHIDAHPLSOTKQAGNDFPIYVAYOATYCAAQAPRPESWDMMC
LIRLPTLHGFTPLVRIAGAVONVETLTHPIPTKTMCMASDLBVLVSTWLVGAVLA
ALAAVCLTGSVAVIVGRIILSGKPAVLPDRVLPQEDMEBCASHLYEQGQULAE
QFOKQALGLIQTATKQAEAPVAVESKRALETFAKMMNFIISGIQYLAGISTLPGN
PAISLMAFTASITSPLTOTTLMENTIIGGVAAOLAPSAASAFAVGAGIAGVSI
GLGKVLVILIAGAGVAGALVAFKWSGEMSPREDIILNLIATLISGALIVGAVCA
ILRHVPGEGAVVMMNRLLIAPSRGNHVSPTHVPSDAAARTQILSLSTITIQULK
RLHOWINBDCSTPCSGSWLNDWDMICVLIDFTWLIQOSKMPRLPGVPLISCGRGK
GVWRGDGIMOTTCPGQAOITGHVKNRIVGPTCSNTWHTGTPINAYTTGPTCPSP

APNYSHALNRVAEEYEVTRVGDFFHYVTGTTDNVNCPCQVAPAEFFTEVDVRLHR
 YAPACKPLREBEVTFQVGLNOYLWGSOLPEPEPEPDVAVLNLSMLTDSHITAEKARL
 AAGSPPLSSSSSSQASABIKATCTTHDSPPADLIENALILRQEMGNTITVBSEN
 KAVILDSFPLRAEBDEBREVSAEILRSRKPFPALPVARIPNPLESKADPDY
 VPPVHGCPLPPTKAPPIPPRRKRTVILRESSVSSALMELATKTSGSSSVYDGT
 ATPPQASDGDKESDVESIMPPLGEPGDPDLSDGMSLRVSEASDVVCCMS
 YMTWGLITPCAESEKLPINPLNSILRHMNVYATSSASLRQKVTFFDLQVLD
 DYHVDLKEMKAKASTYKAKLSIEBAKTLPPHSARSKRGVAKDVRSLSEAVNHI
 RSVMBLLEDETERPIDTTIMAKSEVFCVQPEKGRPARLIVPDIGVRCEGMAIYD
 VSTLPQAVWSSTGFQYSPKORVEFLVNTWSKCPMGFSYTRGCDSTVTSIDRT
 EBSIYOCCLAPARAKAISLTERLYIGGLTVLNSKQONCGRRCRASGLVLTSCGTL
 TCYKATAACRAKALODCTMLVNGDDLVLVCEBAGTQEDASLRVTEAMRTVSAEM
 DPEQPEYDELITSCSNVSVADHDSKRYVYLRDPTPLAAMETARHTVNSML
 GNTIWAYPTLMAMILMTHEFSLIAEOLEKALDOIYACCSISIPLDIPOIIEBLH
 GLSASFHSYSBEINRVASCLRKLVPIPRVRRHARSRAILISOGKARACGKTL
 FMAVVRKTLKLTPIPAASRLDLSGMFVAGISGDIYHSLSRARPRMFMCLLISVGV
 GYLPLNR"

ORIGIN

Query Match 85.9%; Score 889.4; DB 14; Length 9379;
 Best Local Similarity 91.2%; Pred. No. 5.6e-195;
 Matches 944; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 1 ACAGAGTGATGGGTGCGGCTGCAAGGTAGCTCCGCGTGCAAACTCTCTACGG 60
 DB 6696 ACAGAGTGATGGGTGCGGCTGCAAGGTAGCTCCGCGTGCAAACTCTCTACGG 6755
 QY 61 GAGAGGTACATTCAGGTGCGGCTGCAACCAATACCTGTTGGGTGCGAGCTCCATGC 120
 DB 6756 GAGAGGTACATTCAGGTGCGGCTGCAACCAATACCTGTTGGGTGCGAGCTCCATGC 6815
 QY 121 GAGCCGGAACCGGATGAGAGTGTCACTTCATGCTGACCGACCCCTCCCATGACA 180
 DB 6816 GAGCCGGAACCGGATGAGAGTGTCACTTCATGCTGACCGACCCCTCCCATGACA 6875
 QY 181 GCAGAGACGGCTAAGCGCAGCTGCGCAGGAGGCTCTCCCTCTCTGCGCAGCTCTTCA 240
 DB 6876 GCAGAGACGGCTAAGCGCAGCTGCGCAGGAGGCTCTCCCTCTCTGCGCAGCTCTTCA 6935
 QY 241 GCTAGCCAGTTGTCTGGCCCTTCTCTGAGGCGACATATACCCAAATGACTTCCCA 300
 DB 6936 GCTAGCCAGTTGTCTGGCCCTTCTCTGAGGCGACATATACCCCAATGACTTCCCG 6995
 QY 301 GAGGCTGACTCATTCAGGCGCAACTCTCTGCGCGCATGAGATGGCGGGGACATTAAC 360
 DB 6996 GAGGCTGACTCATTCAGGCGCAACTCTCTGCGCGCATGAGATGGCGGGGACATTAAC 7055
 QY 361 CGCGTGAAGTCAGAGAAAGAGTAGTATCTCTGGAATCTTTCGACCCGCTCCGAGCGGAG 420
 DB 7056 CGCGTGAAGTCAGAGAAAGAGTAGTATCTTGGACTCTTTCGACCCGCTTTCGAGCGGAG 7115
 QY 421 GAGGATGAGCGGGAAGTGTCCGTCCGCGCGAGATCTCTGCGGAAATCCAGAAATTCACA 480
 DB 7116 GAGGATGAGCGGGAAGTGTCCGTCCGCGCGAGATCTCTGCGGAAATCCAGAAATTCACA 7175
 QY 481 CCAGCGATGCGCGCATGCGGCAAGCGCGGATTAACAACCTCTCCGCTGCTGAGTCTGGAAG 540
 DB 7176 CCAGCGCTGCGCATGCGGCAAGCGCGGATTAACAACCTCTCCGCTGCTGAGTCTGGAAG 7235
 QY 541 GCCCGGGAAGTACGTCCTCCAGTGTGTCATGCGGTCGCCATGCGCACTTACAAAGCCCT 600
 DB 7236 GCCCGGGAAGTACGTCCTCCAGTGTGTCATGCGGTCGCCATGCGCACTTACAAAGCCCT 7295
 QY 601 CCTATAACAACCTTCAAGGAGAAAGAGACAGTTGTTCTGACAGATCCACCGTGTCTTCT 660
 DB 7296 CCTATAACAACCTTCAAGGAGAAAGAGACAGTTGTTCTGACAGATCCACCGTGTCTTCT 7355
 QY 661 GCCCTGGCGGAGCTTGGCACAAGAGCTTTTGTAGCTCCGGAACGTCGCGCGTTCGACAGC 720
 DB 7356 GCCCTGGCGGAGCTTGGCACAAGAGCTTTTGTAGCTCCGGAACGTCGCGCGTTCGACAGC 7415
 QY 721 GGCACGGCAACGCGCCCTCTGACCAATCTCCGACGACGCGGAGCAGAGATCTGACGTT 780

DB 7416 GGCACGGCAACTGCTCCCTCCGTCACAGGCTCCGACGAGCGGCAAGAAATCCGACGTT 7475
 QY 781 GACTCGATATCTCCCATAGCCCCCTTGAAGGGGAGCCCGGGGAGCCCGCATCTCAGCGAC 840
 DB 7476 GACTCGATATCTCCCATAGCCCCCTTGAAGGGGAGCCCGGGGAGCCCGCATCTCAGCGAC 7535
 QY 841 GGGCTTTGGTCTACCGCTGAGTGAAGAGCGCGTGAAGACGTCGCTGCTGCTGATGTC 900
 DB 7536 GGGTCTTGGTCTACCGCTGAGTGAAGAGCGCGTGAAGACGTCGCTGCTGCTGATGTC 7595
 QY 901 TACACATGACAGCGGCTCTGATCAGCGCATGCGCTGCGAGGAAAGCAAGTCCCATC 960
 DB 7596 TACACATGACAGCGGCTCTGATCAGCGCATGCGCTGCGAGGAAAGCAAGTCCCATC 7655
 QY 961 AAGCGTTGAGCAACTTTGCTGCTGCTGCAACCAATAGTCTACGCTACCATCCCGC 1020
 DB 7656 AAGCGTTGAGCAACTTTGCTGCTGCTGCAACCAATAGTCTACGCTACCATCCCGC 7715
 QY 1021 AGCGCAAGCCAGCGG 1035
 DB 7716 AGCGCAAGCCAGCGG 7730

Search completed: February 19, 2005, 07:08:43
 Job time : 4603.2 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2005, 20:55:25 ; Search time 579.412 Seconds
(without alignments)
10574.404 Million cell updates/sec

Title: US-09-664-363-4

Perfect score: 1035
Sequence: 1 ACAGAGATGATGATGCGGTCG.....CCCGACGCCAGCAGCCG 1035

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1035	100.0	2790	2	AAQ46191	AaQ46191 PT-NANBH
2	1035	100.0	3372	2	AAQ46202	AaQ46202 PT-NANBH
3	1033.4	99.8	1035	2	AAQ12237	AaQ12237 Clone UG3
4	1033.4	99.8	1788	2	AAQ46193	AaQ46193 PT-NANBH
5	1025.4	99.1	3750	2	AAQ12241	AaQ12241 Encodes P
6	1025.4	99.1	7065	2	AAQ46195	AaQ46195 PT-NANBH
7	889.4	85.9	7989	10	ADD93733	AbK85773 Hepatitis
8	885.8	85.6	8642	6	ABK8574	AbK8574 Hepatitis
9	884.6	85.5	7987	6	AAD25321	Aad25321 Hepatitis
10	884.6	85.5	7989	3	AAA8968	Aaa8968 Hepatitis
11	884.6	85.5	7989	6	AAD25322	Aad25322 Hepatitis
12	884.6	85.5	7989	12	ADJ57845	Adj57845 HCV repli
13	884.6	85.5	7991	6	AA147279	Aa147279 Hepatitis
14	884.6	85.5	7992	6	AA147276	Aa147276 Hepatitis
15	884.6	85.5	7992	6	AA147280	Aa147280 Hepatitis
16	884.6	85.5	8001	3	AAA8967	Aaa8967 Hepatitis
17	884.6	85.5	8637	3	AAA8966	Aaa8966 Hepatitis
18	884.6	85.5	8638	6	ABK8595	AbK8595 Hepatitis
19	884.6	85.5	8638	6	ABK8577	AbK8577 Hepatitis
20	884.6	85.5	8639	6	ABK8573	AbK8573 Hepatitis

ALIGNMENTS

21	884.6	85.5	8649	3	AAA8969	Aaa8969 Hepatitis
22	884.6	85.5	9605	6	ABK91424	AbK91424 Hepatitis
23	884.6	85.5	9605	6	ABK91411	AbK91411 Hepatitis
24	884.6	85.5	9605	6	ABK91428	AbK91428 Hepatitis
25	884.6	85.5	9605	6	ABK91425	AbK91425 Hepatitis
26	884.6	85.5	9605	6	ABK91426	AbK91426 Hepatitis
27	884.6	85.5	9608	6	ABK91427	AbK91427 Hepatitis
28	884.6	85.5	10690	6	ABK91243	AbK91243 Hepatitis
29	884.6	85.5	10690	6	ABK91242	AbK91242 Hepatitis
30	884.6	85.5	10690	6	ABK91412	AbK91412 Hepatitis
31	884.6	85.5	10690	9	ACA61697	AcA61697 Hepatitis
32	884.6	85.5	10690	10	ADC83762	AdC83762 PHCNeo17
33	884.6	85.5	10691	6	ABK91423	AbK91423 Hepatitis
34	884.6	85.5	10693	6	ABK91443	AbK91443 Hepatitis
35	884.6	85.5	10693	6	ABK91438	AbK91438 Hepatitis
36	884.6	85.5	11076	3	AAA8965	Aaa8965 Hepatitis
37	884.6	85.5	11313	12	ADP86273	AdP86273 Hepatitis
38	884.6	85.5	11313	12	ADP86264	AdP86264 Hepatitis
39	884.6	85.5	11313	12	ADP86271	AdP86271 Hepatitis
40	884.6	85.5	11313	12	ADP86275	AdP86275 Hepatitis
41	883	85.3	7141	6	AAD25333	Aad25333 Hepatitis
42	883	85.3	7789	6	AAD25350	Aad25350 Hepatitis
43	883	85.3	7980	10	ADD93724	Add93724 Hepatitis
44	883	85.3	7980	10	ADD93725	Add93725 Hepatitis
45	883	85.3	7987	6	AAD25329	Aad25329 Hepatitis

RESULT 1

ID	AAQ46191	standard; DNA; 2790 BP.
AC	AAQ46191;	
XX		
DT	27-AUG-2003	(revised)
DT	25-MAR-2003	(revised)
DT	24-FEB-1994	(first entry)
XX		
DE	PT-NANBH virus BHC-11 fusion protein.	
XX		
KM	Parenterally transmitted non A non B hepatitis; PT-NANBH;	
KM	hepatitis C virus; HCV; NS5; E1; linker; detection; diagnosis; antigen;	
KW	vaccine; BHC-11; replicase; core protein;	
KW	Antograpa californica nuclear polydrosis virus; AcNPV; polyhedrin; ss.	
XX		
OS	Hepatitis virus.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	1..2790
FT		/*tag= a
FT		/product= "fusion_protein"
FT		1..63
FT		/*tag= b
FT		/label= AcNPV_polyhedrin_N-terminal
FT		64..1851
FT		/*tag= c
FT		/label= PT-NANBH_NS5
FT		1852..1875
FT		/*tag= d
FT		/note= "Synthetic linker region"
FT		1876..2706
FT		/*tag= e
FT		/label= PT-NANBH_core_and_E1_regions
FT		2707..2790
FT		/*tag= f
FT		/note= "Polyhedrin gene sequence read out-of-frame"
PN	WO9317110-A2.	
XX		
PD	02-SEP-1993.	
XX		

PF 19-FEB-1993; 93WO-GB000345.
 XX 21-FEB-1992; 92GB-00003803.
 XX (WELL) WELLCOME FOUND LTD.
 PA Parker D, Rodgers BC;
 PI WPI; 1993-288415/36.
 DR P-PSDB; AAR41431.
 XX
 PT New recombinant polypeptide for diagnosing hepatitis C - contains three
 PT distinct antigens from different viral regions, also useful in protective
 PT vaccine.
 PS
 XX Disclosure: Page 20-24; 99pp; English.
 CC GB-A-2239245 discloses a recombinant polypeptide BHC-11 which comprises
 CC an antigen obtained from the non-structural coding region (NS) (the 3'
 CC end) and one antigen from the structural coding region (S) (the 5' end)
 CC of the NANBH virus. Specifically BHC-11 (AA046191) contains a portion of
 CC the non-structural region of the virus, called NS5, (putative replicase)
 CC at the N-terminus joined via a synthetic linker to a portion of the
 CC structural region which contains almost all the core protein sequence (9
 CC amino acids from the N-terminal are not present) and a part of a sequence
 CC from the structural region called E1. It is disclosed that BHC-11 may be
 CC used in diagnosis of PT-NANBH. If at least three different PT-NANBH
 CC antigens are used to screen for PT-NANBH, the screening is much more
 CC sensitive as compared to the use of only two PT-NANBH antigens. Pref.
 CC antigens are described in AA046192-94. Two new antigenic regions of the
 CC PT-NANBH genome are given in AA046198-99. AA046202 describes an improved
 CC PT-NANBH recombinant polypeptide. (Updated on 25-MAR-2003 to correct PN
 CC field.) (Updated on 27-AUG-2003 to correct OS field.)
 XX
 SQ Sequence 2790 BP; 585 A; 861 C; 789 G; 555 T; 0 U; 0 Other;

Query Match 100.0%; Score 1035; DB 2; Length 2790;
 Best Local Similarity 100.0%; Pred. No. 2.6e-268;
 Matches 1035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGAGAGTGGATGGGTGGCTGCAAGTACGCTCCGGCTGCAAACTCTCTACGG 60
 DB 64 AAGAGAGTGGATGGGTGGCTGCAAGTACGCTCCGGCTGCAAACTCTCTACGG 123
 QY 61 GAGAGAGTCAATTCAGGTCGGGCTCAACCAATACCTGTTGGGTGCGAGCTCCATGC 120
 DB 124 GAGAGAGTCAATTCAGGTCGGGCTCAACCAATACCTGTTGGGTGCGAGCTCCATGC 183
 QY 121 GAGCCCGAACCAGATGTAGCACTGCTCACTTCATGCTCAGCCGACCCCTCCCATCA 180
 DB 184 GAGCCCGAACCAGATGTAGCACTGCTCACTTCATGCTCAGCCGACCCCTCCCATCA 243
 QY 181 GAGAGAGCGCTTAACGCGAGGCTGCGCCAGGGGCTCTCCCTCTCTGGCCAGCTTTCA 240
 DB 244 GAGAGAGCGCTTAACGCGAGGCTGCGCCAGGGGCTCTCCCTCTCTGGCCAGCTTTCA 303
 QY 241 GCTAACCAATTGTCTGCGCTTCTCTGAAAGCGACATACATTACCAAAATGATTTCCA 300
 DB 304 GCTAACCAATTGTCTGCGCTTCTCTGAAAGCGACATACATTACCAAAATGATTTCCA 363
 QY 301 GAGCGTGAATCATGAGGCGCAACTCTGTGGCGGATGAGATGGGCGGAGCATTAAC 360
 DB 364 GAGCGTGAATCATGAGGCGCAACTCTGTGGCGGATGAGATGGGCGGAGCATTAAC 423
 QY 361 GCGGTGAGTCAAGAAAGTAAATCTTGAAGTCTTTGACCCGCTCCGAGCGAG 420
 DB 424 GCGGTGAGTCAAGAAAGTAAATCTTGAAGTCTTTGACCCGCTCCGAGCGAG 483
 QY 421 GAGGATGAGCGGGAAGTGTCCGTCCGCGGAGATCTCTGCGAAATTCAGAAATTTCCA 480
 DB 484 GAGGATGAGCGGGAAGTGTCCGTCCGCGGAGATCTCTGCGAAATTCAGAAATTTCCA 543
 QY 481 CAGGATGAGCGGGAAGTGTCCGTCCGCGGATTAACAACCTCTGCTGAGTCTGGAAG 540

DB 544 CAGGATGAGCGGGAAGTGTCCGTCCGCGGATTAACAACCTCTGCTGAGTCTGGAAG 603
 QY 541 GCGCGGAGTCAATTCAGGTCGGGCTCAACCAATACCTGTTGGGTGCGAGCTCCATGC 600
 DB 604 GCGCGGAGTCAATTCAGGTCGGGCTCAACCAATACCTGTTGGGTGCGAGCTCCATGC 663
 QY 601 CTTATACCACTCCACGAGAAAGAGACAGTGTCTTGAAGATTCACCGGTGCTTCT 660
 DB 664 CTTATACCACTCCACGAGAAAGAGACAGTGTCTTGAAGATTCACCGGTGCTTCT 723
 QY 661 GCGTGGCGGAGCTTTGCCAACAAGCTTTGTGATGCTCCGACCGGTGCGGTGACAGC 720
 DB 724 GCGTGGCGGAGCTTTGCCAACAAGCTTTGTGATGCTCCGACCGGTGCGGTGACAGC 783
 QY 721 GGCAGGGAACCGCCCTCTGACCAATCTCCACGACGCGGAGAGATTCGACCTT 780
 DB 784 GGCAGGGAACCGCCCTCTGACCAATCTCCACGACGCGGAGAGATTCGACCTT 843
 QY 781 GAGTGTATTCTCCATGCCCCCTTGAAGGAGGCGGAGGAGCCCGATCTCAGCGAC 840
 DB 844 GAGTGTATTCTCCATGCCCCCTTGAAGGAGGCGGAGGAGCCCGATCTCAGCGAC 903
 QY 841 GGTCTTGTATCCGTGAGTGAAGAGCGGAGGAGCGTCTGCTGCTGATGTC 900
 DB 904 GGTCTTGTATCCGTGAGTGAAGAGCGGAGGAGCGTCTGCTGCTGATGTC 963
 QY 901 TACACATGAGAGCGCTCTGATCAACGCAATGCGTGGGAGAAAGCACTGCGCATC 960
 DB 964 TACACATGAGAGCGCTCTGATCAACGCAATGCGTGGGAGAAAGCACTGCGCATC 1023
 QY 961 AACGCTTGAACCACTTTGCTGCTGCTGCAACCAATGCTTACGATACCATCCCGC 1020
 DB 1024 AACGCTTGAACCACTTTGCTGCTGCTGCAACCAATGCTTACGATACCATCCCGC 1083
 QY 1021 AGCGCAAGCCAGCGG 1035
 DB 1084 AGCGCAAGCCAGCGG 1098

RESULT 2
 AA046202
 ID AA046202 standard; cDNA to mRNA; 3372 BP.
 XX
 AC AA046202;
 XX
 DT 27-AUG-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 24-FEB-1994 (first entry)
 XX
 DE PT-NANBH NS-NS3-core recombinant polypeptide.
 XX
 KM Parenterally transmitted non A non B hepatitis; PT-NANBH;
 KM hepatitis C virus; HCV; NS3; NS5; core; E1; linker; detection; diagnosis;
 KM antigen; vaccine; BHC-28; replicase; core protein;
 KM Autographa californica nuclear polyhedrosis virus; AcNPV; polyhedrin; ss.
 OS Hepatitis virus.
 XX
 FH Key
 FH CDS 1. 3372
 FT /tag= a
 FT /product= "fusion_protein"
 FT misc_feature 1. 63
 FT /tag= b
 FT /label= AcNPV_polyhedrin_N-terminal
 FT misc_feature 64. 1852
 FT /tag= c
 FT /label= PT-NANBH_NS5
 FT misc_feature 1853. 1858
 FT /tag= d
 FT /note= "PstI restriction site"
 FT misc_feature 1859. 2434

```

FT      /*tag= e
FT      /label= PT-NANBH_NS3
FT      misc_feature
FT      2435..2457
FT      /*tag= f
FT      /note= "Paxi restriction site"
FT      misc_feature
FT      2458..3288
FT      /*tag= g
FT      /label= PT-NANBH_core_and_E1_regions
FT      misc_feature
FT      3289..3372
FT      /*tag= h
FT      /note= "Polyhedrin gene sequence read out-of-frame"
PN      MO9317110-A2.
PD      02-SEP-1993.
XX
XX      19-FEB-1993; 93WO-GB000345.
XX
XX      21-FEB-1992; 92GB-00003803.
XX
XX      (WELL ) WELLCOME FOUND LTD.
XX
XX      Parker D, Rodgers BC;
XX
XX      WPI; 1993-288415/36.
XX      P-PSDB; AAR41439.
XX
XX      New recombinant polypeptide for diagnosing hepatitis C - contains three
XX      distinct antigens from different viral regions, also useful in protective
XX      vaccines.
XX
XX      Claim 7; Page 85-90; 99pp; English.
XX
XX      The NS3 specific region of pDX200 was amplified by PCR using primers D360
XX      and D361 (AAQ46201 and AAQ54882). The obtained fragment was then cloned
XX      into pDX136. Transformants were analysed to identify those which
XX      contained the NS3 sequence inserted in the correct orientation between
XX      the NS5 and core parts of pDX136, this was called pDX208 (AAQ46202).
XX      Recombinant baculovirus BHC-28 was produced. Insect cells infected with
XX      BHC-28 produce antigen NS5-NS3-core. If at least three different PT-NANBH
XX      antigens are used to screen for PT-NANBH, the screening is much more
XX      sensitive as compared to the use of only two PT-NANBH antigens. Pref.
XX      CC antigens are described in AAQ46192-94. Two new antigenic regions of the
XX      PT-NANBH genome are given in AAQ46198-99. AAQ46202 describes an improved
XX      PT-NANBH recombinant polypeptide. (Updated on 25-MAR-2003 to correct PN
XX      field.) (Updated on 27-AUG-2003 to correct OS field.)
SQ      Sequence 3372 BP; 714 A; 1026 C; 954 G; 678 T; 0 U; 0 Other;

Query Match      100.0%; Score 1035; DB 2; Length 3372;
Best Local Similarity 100.0%; Pred. No. 2.8e-268;
Matches 1035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      301 GACGCTGACCTCATGAGGCGCAACCTCTGTGGCGGCAAGAGATGGGCGGAGCATTTACC 360
DB      |||
DB      364 GAGCTGACCTCATGAGGCGCAACCTCTGTGGCGGCAAGAGATGGGCGGAGCATTTACC 423
QY      361 CGGCTGAGTCAAGAGAACAGAGTAGTAATCTTGAATCTTTTGCACCCGCTCGAGCGAG 420
DB      |||
DB      424 CGGCTGAGTCAAGAGAACAGAGTAGTAATCTTGAATCTTTTGCACCCGCTCGAGCGAG 483
QY      421 GAGGATGAGCGGAGAGTGTCCGTCGGCGGAGATCTCTGCGGAATCCAGAAATTTCCCA 480
DB      |||
DB      484 GAGGATGAGCGGAGAGTGTCCGTCGGCGGAGATCTCTGCGGAATCCAGAAATTTCCCA 543
QY      481 CCAGGATGCGCGGCAATGGGCAAGCGCCGGATTCAACCTCCGCTGTGAGTCTTGAAG 540
DB      |||
DB      544 CCAGGATGCGCGGCAATGGGCAAGCGCCGGATTCAACCTCCGCTGTGAGTCTTGAAG 603
QY      541 GCCCGGAGCTAGTCCCTCCAGTGTATCATGGGTGCGCACTGCACTTAAGACCCCT 600
DB      |||
DB      604 GCCCGGAGCTAGTCCCTCCAGTGTATCATGGGTGCGCACTGCACTTAAGACCCCT 663
QY      601 CCTATTCACCTTCACAGAGAAAGAGACAGTTTGTTCACAGATCCACCGTCTTCT 660
DB      |||
DB      664 CCTATTCACCTTCACAGAGAAAGAGACAGTTTGTTCACAGATCCACCGTCTTCT 723
QY      661 GCCCTGGCGGAGCTTGCACAAAGGCTTTGGTAGCTCCGACCGGCGGCGTGCAGCAGC 720
DB      |||
DB      724 GCCCTGGCGGAGCTTGCACAAAGGCTTTGGTAGCTCCGACCGGCGGCGTGCAGCAGC 783
QY      721 GGCACGGCAACCGCCCTCTCTGACCAATCTCTCGACGAGCGGAGACAGATCTGACGTT 780
DB      |||
DB      784 GGCACGGCAACCGCCCTCTCTGACCAATCTCTCGACGAGCGGAGACAGATCTGACGTT 843
QY      781 GAGTCGTATTCTCCATGCCCCCTTTGAGGGGAGACCGGGGACCCGATCTCAGCGAC 840
DB      |||
DB      844 GAGTCGTATTCTCCATGCCCCCTTTGAGGGGAGACCGGGGACCCGATCTCAGCGAC 903
QY      841 GGGTCTGGTCTACCGGTGAGAGAGCGCGGTGAGAGAGTGTGCTGCTGATGTC 900
DB      |||
DB      904 GGGTCTGGTCTACCGGTGAGAGAGCGCGGTGAGAGAGTGTGCTGCTGATGTC 963
QY      901 TACACATGACAGGCGCTCTGATCAAGCATGCGTCTCGAGAGAAAGCAAGCTGCCATC 960
DB      |||
DB      964 TACACATGACAGGCGCTCTGATCAAGCATGCGTCTCGAGAGAAAGCAAGCTGCCATC 1023
QY      961 AACGCGTTGAGCAACTTTTGTGCGTCAACCAACATGTTTACGTAACCAATCCCGC 1020
DB      |||
DB      1024 AACGCGTTGAGCAACTTTTGTGCGTCAACCAACATGTTTACGTAACCAATCCCGC 1083
QY      1021 AGCGCAAGCGACGCG 1035
DB      |||
DB      1084 AGCGCAAGCGACGCG 1098

RESULT 3
AAQ12237
ID      AAQ12237 standard; DNA; 1035 BP.
AC      AAQ12237;
XX
AC      25-MAR-2003 (revised)
DT      06-SEP-1991 (first entry)
XX
DE      Clone U33 encoding PT-NANBH virus antigenic portion.
XX
XX      post-translational non-A, non-B hepatitis; virus; vaccine; ss.
XX
XX      Non-A.
OS      non-B hepatitis virus.
XX
XX      GB2239245-A.
XX
XX      26-JUN-1991.
XX

```

PF 17-DEC-1990; 90GB-00027250.
XX 18-DEC-1989; 89GB-00028562.
XX 27-FEB-1990; 90GB-00004414.
PR 03-MAR-1990; 90GB-00004814.
XX
PA (WELL) WELLCOME FOUNDT LTD.
PA (HIGH/) HIGHFIELD P E.
XX
PI Highfield PE, Rodgers BC, Tedder RS, Barbara JAJ;
XX
DR WPI, 1991-187584/26.
DR P-PSDB; AAR12595.
XX
PT Post-transferritional non-A non-B hepatitis poly:peptide(s) - and also DNA
PT and antibodies used in diagnostic assays and in vaccines.
XX
PS Claim 10; Page 53-55; 108pp; English.
XX
CC This sequence is a non-structural region of the PT-NANBH viral genome
CC encoding an antigenic polypeptide. It was isolated from serum of humans
CC infected by the virus. Genomic RNA from viruses pelleted from the
CC infected serum was used to prepare a cDNA library in lambda gt11 which
CC was screened with antibodies from the original serum and oligonucleotide
CC probes. One of the positive plaques to be obtained was J63. DNA was
CC extracted from denatured phage stocks, amplified by PCR and sequenced.
CC See also AAQ12236 and AAQ12238-Q12242. (Updated on 25-MAR-2003 to correct
CC PA field.)
XX
SQ Sequence 1035 BP, 212 A, 340 C, 296 G, 187 T, 0 U, 0 Other;

Query Match 99.8%; Score 1033.4; DB 2; Length 1035;
Best Local Similarity 99.9%; Pred. No. 5.2e-268;
Matches 1034; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAGAGAGTGGATGGGCTGGCTGCAAGGTACGCTCCGGCGTCAAACTCTCTACGG 60
DB 1 AAGAGAGTGGATGGGCTGGCTGCAAGGTACGCTCCGGCGTCAAACTCTCTACGG 60
QY 61 GAGGAGGTGACATTCAGGTCGGGGCTCAACCAATACCTGGTGGGTCGAGTCCATGC 120
DB 61 GAGGAGGTGACATTCAGGTCGGGGCTCAACCAATACCTGGTGGGTCGAGTCCATGC 120
QY 121 GAGCCCGAACCAGATGTAGCACTGCTCACTTCATGCTCAGACCCCTCCACATACA 180
DB 121 GAGTCCGAACCGGATGTAGCACTGCTCACTTCATGCTCAGACCCCTCCACATACA 180
QY 181 GCAAGACCGCTTAAGCGCAGGCTGCGCAGGGGGTCTCCCTCTTGGCCAGCTCTTCA 240
DB 181 GCAAGACCGCTTAAGCGCAGGCTGCGCAGGGGGTCTCCCTCTTGGCCAGCTCTTCA 240
QY 241 GCTAGCCAGTTGTCTGGCCCTCTCGAAGGGGAGCATATATACCAAAATACCTTCCA 300
DB 241 GCTAGCCAGTTGTCTGGCCCTCTCGAAGGGGAGCATATATACCAAAATACCTTCCA 300
QY 301 GACGCTGACCTCATGAGGCAACCTCTGTGGCGGATGAGATGGGCGGGACATTACC 360
DB 301 GACGCTGACCTCATGAGGCAACCTCTGTGGCGGATGAGATGGGCGGGACATTACC 360
QY 361 CGCGTGAAGTCAAGAAACAAGTATATCTTGAAGCTTTTGAACCCGCTCCGAGGAG 420
DB 361 CGCGTGAAGTCAAGAAACAAGTATATCTTGAAGCTTTTGAACCCGCTCCGAGGAG 420
QY 421 GAGGATGAGCGGGAAGTGTCCGTCGCGGAGGATCTCGGGAATCCAAAGAAATTTCCA 480
DB 421 GAGGATGAGCGGGAAGTGTCCGTCGCGGAGGATCTCGGGAATCCAAAGAAATTTCCA 480
QY 481 CCAGGATGCTCCGATGGGACAGCCCGATTAACAACCTCCCTGCTGAGTCTGGAAG 540
DB 481 CCAGGATGCTCCGATGGGACAGCCCGATTAACAACCTCCCTGCTGAGTCTGGAAG 540
QY 541 GCCCGGACTACGTCCCTCAGTGGTACATGGGTGCCCATCTGACCTTAAGAACCCT 600
DB 541 GCCCGGACTACGTCCCTCAGTGGTACATGGGTGCCCATCTGACCTTAAGAACCCT 600

DB 541 GCCCGGACTACGTCCCTCAGTGGTACATGGGTGCCCATCTGACCTTAAGAACCCT 600
QY CCTATACCACTCCACGGAGAAAGAGCAGTTGTTGACAGAAATCCACCGTGTCTCT 660
DB CCTATACCACTCCACGGAGAAAGAGCAGTTGTTGACAGAAATCCACCGTGTCTCT 660
QY 661 GCCCTGGGAGCTTGGCACAAGGCTTTTGTAGCTTCGGACCGTCCGCGACAGC 720
DB 661 GCCCTGGGAGCTTGGCACAAGGCTTTTGTAGCTTCGGACCGTCCGCGACAGC 720
QY 721 GGCACGGACACCGCCCTCTCTGACCAATCTTCGACGACGGCGGACGATCTGAGCTT 780
DB 721 GGCACGGACACCGCCCTCTCTGACCAATCTTCGACGACGGCGGACGATCTGAGCTT 780
QY 781 GAGTGTATTCCTCCATGCCCCCTTGAAGGGAGACCGGGGACCCCGATCTGAGCAG 840
DB 781 GAGTGTATTCCTCCATGCCCCCTTGAAGGGAGACCGGGGACCCCGATCTGAGCAG 840
QY 841 GGGTCTTGCTTACCGTAGTAGAGGCGGAGAGCGTGTCTGCTGATGTC 900
DB 841 GGGTCTTGCTTACCGTAGTAGAGGCGGAGAGCGTGTCTGCTGATGTC 900
QY 901 TACATATGACAGGCGCTCTGATCACCGCATGCGTGGGAGAAAGCAAGTCCCATC 960
DB 901 TACATATGACAGGCGCTCTGATCACCGCATGCGTGGGAGAAAGCAAGTCCCATC 960
QY 961 AACGGTTGAGCAACTTTTGTGCTGACCAACAATGCTACGCTACCAATCCCGC 1020
DB 961 AACGGTTGAGCAACTTTTGTGCTGACCAACAATGCTACGCTACCAATCCCGC 1020
QY 1021 AGCGCAAGCCAGCG 1035
DB 1021 AGCGCAAGCCAGCG 1035
RESULT 4
AAQ46193
ID AAQ46193 standard; DNA; 1788 BP.
XX
AC AAQ46193;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 24-FEB-1994 (first entry)
XX
DE PT-NANBH virus NS5 protein.
XX
KW Parenterally transmitted non A non B hepatitis; PT-NANBH;
KW hepatitis C virus; HCV; detection; diagnosis; antigen; vaccine; ss.
OS Hepatitis virus.
XX
PN W09317110-A2.
XX
PD 02-SEP-1993.
XX
PF 19-FEB-1993; 93MO-GB000345.
XX
PR 21-FEB-1992; 92GB-00003803.
XX
PA (WELL) WELLCOME FOUNDT LTD.
XX
PI Parker D, Rodgers BC;
XX
DR WPI, 1993-288415/36.
DR P-PSDB; AAR41433.
XX
PT New recombinant polypeptide for diagnosing hepatitis C - contains three
PT distinct antigens from different viral regions, also useful in protective
XX vaccines.
XX
PS Claim 7; Page 33-36; 99pp; English.
XX

CC If at least three different PT-NANBH antigens are used to screen for PT-
 CC NANBH, the screening is much more sensitive as compared to the use of
 CC only two PT-NANBH antigens. Pref. antigens are described in AAQ46192-94.
 CC Two new antigenic regions of the PT-NANBH genome are given in AAQ46198-
 CC 99. AAQ46202 describes an improved PT-NANBH recombinant polypeptide.
 CC (Updated on 25-MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to
 CC correct OS field.)

XX Sequence 1788 BP; 401 A; 544 C; 498 G; 345 T; 0 U; 0 Other;

Query Match 99.8%; Score 1033.4; DB 2; Length 1788;

Best Local Similarity 99.9%; Pred. No. 6.2e-268;

Matches 1034; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ACAGAAAGTGGATGGGCTGGCTGCAAGGTACGCTCCGGCGTCAAACTCTCTTACCG 60
DB 1 ACAGAAAGTGGATGGGCTGGCTGCAAGGTACGCTCCGGCGTCAAACTCTCTTACCG 60
QY 61 GAGGAGGTGCAATTCAGGTCGGGCTCAACCAATACCTGGTGGGTGCGACGTCCTCCATGC 120
DB 61 GAGGAGGTGCAATTCAGGTCGGGCTCAACCAATACCTGGTGGGTGCGACGTCCTCCATGC 120
QY 121 GAGCCCGAACCAGATAGCAAGTGTCACTTCATGCTGCAACGCCCTCCCATCACA 180
DB 121 GAGCCCGAACCAGATAGCAAGTGTCACTTCATGCTGCAACGCCCTCCCATCACA 180
QY 181 GCAAGACGCGCTAAGCGCAGGCTGCGCAAGGGGTCTCCCGCTCTTGGCCAGCTCTTCA 240
DB 181 GCAAGACGCGCTAAGCGCAGGCTGCGCAAGGGGTCTCCCGCTCTTGGCCAGCTCTTCA 240
QY 241 GGTAGCAATGTCTGCGCCCTCTCTGAAAGCGCAATACCTTCCCAAAATGACTTCCCA 300
DB 241 GGTAGCAATGTCTGCGCCCTCTCTGAAAGCGCAATACCTTCCCAAAATGACTTCCCA 300
QY 301 GACGCTGACTATCGAGGCAACCTCTGTGCGCGCATGAGATGGGCGGAGCATTAAC 360
DB 301 GACGCTGACTATCGAGGCAACCTCTGTGCGCGCATGAGATGGGCGGAGCATTAAC 360
QY 361 GCGGTGAGTCAAGAAACAAGTATCTCTGACTCTTTGCAACCGCTCCGAGCGGAG 420
DB 361 GCGGTGAGTCAAGAAACAAGTATCTCTGACTCTTTGCAACCGCTCCGAGCGGAG 420
QY 421 GAGGATGAGGGGGAAGTCCCGTCCGGGGGAGATCTCGGGAATCCAAAGAAATCCCA 480
DB 421 GAGGATGAGGGGGAAGTCCCGTCCGGGGGAGATCTCGGGAATCCAAAGAAATCCCA 480
QY 481 CGAGCATGCCGATGAGGACAGCCCGGATTAACAACCTCCGTCTGAGTCTGGAAG 540
DB 481 CGAGCATGCCGATGAGGACAGCCCGGATTAACAACCTCCGTCTGAGTCTGGAAG 540
QY 541 GCCCGGACTACGTCCTCCAGTGTACATGGGTGCCCATCTGCACTTAAGACCCCT 600
DB 541 GCCCGGACTACGTCCTCCAGTGTACATGGGTGCCCATCTGCACTTAAGACCCCT 600
QY 601 CCTATACCACTCCAGGGAAGAAAGAGACAGTGTCTGACAAATCCACCGGTCTTCT 660
DB 601 CCTATACCACTCCAGGGAAGAAAGAGACAGTGTCTGACAAATCCACCGGTCTTCT 660
QY 661 GCCCTGCGGAGCTTCCCAAAAGCTTTGTGAGTCTCGGACCGTCCGCGCTGACAGC 720
DB 661 GCCCTGCGGAGCTTCCCAAAAGCTTTGTGAGTCTCGGACCGTCCGCGCTGACAGC 720
QY 721 GGCACGGCAACCGCCCTCTCTGACCAATCTTCGACAGCGCGGACAGATGTGACGT 780
DB 721 GGCACGGCAACCGCCCTCTCTGACCAATCTTCGACAGCGCGGACAGATGTGACGT 780
QY 781 GAGTGTATTTCTTCATGCCCCCTTGAAGGGGGAACCCCGATCTCAACGAC 840
DB 781 GAGTGTATTTCTTCATGCCCCCTTGAAGGGGGAACCCCGATCTCAACGAC 840
QY 841 GGGTCTTGTCTACCGTGAAGAGAGCGGATGAGAGCGTCTGCTGCTGATGTCC 900
DB 841 GGGTCTTGTCTACCGTGAAGAGAGCGGATGAGAGCGTCTGCTGCTGATGTCC 900

```

```

QY 901 TACACATGACAGCGCTGTGATCAAGCATGCGCTGCGGAGGAAGCAAGCTGCCATC 960
DB 901 TACACATGACAGCGCTGTGATCAAGCATGCGCTGCGGAGGAAGCAAGCTGCCATC 960
QY 961 AACGCGTGAACAATCTTTGCTGCGTCAACCAACATGATCTACGCTACATCCGC 1020
DB 961 AACGCGTGAACAATCTTTGCTGCGTCAACCAACATGATCTACGCTACATCCGC 1020
QY 1021 ACGGCAAGCCAGCGG 1035
DB 1021 ACGGCAAGCCAGCGG 1035

```

RESULT 5

```

AAQ12241
ID AAQ12241 standard; cDNA; 3750 BP.
XX
AC AAQ12241;
XX
DT 25-MAR-2003 (revised)
DT 17-SEP-1991 (first entry)
XX
DE Encodes portion of PT-NANBH viral non-structural protein.
XX
KW post-transfusional non-A, non-B hepatitis; virus; vaccine; ss.
XX
OS Non-A.
OS non-B hepatitis virus.
XX
PN GB239245-A.
XX
PD 26-JUN-1991.
XX
PF 17-DEC-1990; 90GB-00027250.
XX
PR 18-DEC-1989; 89GB-00028562.
PR 27-FEB-1990; 90GB-00004414.
PR 03-MAR-1990; 90GB-00004814.
XX
PA (WELL) WELLCOME FOUND LTD.
PA (HIGH) HIGHFIELD P E.
XX
PI Highfield PE, Rodgers BC, Tedder RS, Barbara JAJ;
XX
DR WPI; 1991-187584/26.
DR P-PSDB; AAR12599.
XX
PT Post-transfusional non-A non-B hepatitis poly:peptide(s) - and also DNA
PT and antibodies used in diagnostic assays and in vaccines.
XX
PS Claim 10; Page 88-97; 108pp; English.
XX
CC This sequence probably encodes viral non-structural proteins of the PT-
CC NANBH viral genome which are antigenic. It was isolated from serum of
CC humans infected by the virus. See also AAQ12236-40 and AAQ12242. (Updated
CC on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 3750 BP; 794 A; 1140 C; 1072 G; 744 T; 0 U; 0 Other;

```

Query Match 99.1%; Score 1025.4; DB 2; Length 3750;
 Best Local Similarity 99.4%; Pred. No. 1.1e-265;
 Matches 1029; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

QY 1 ACAGAAAGTGGATGGGCTGGCTGCAAGGTACGCTCCGGCGTCAAACTCTCTTACCG 60
DB 1696 ACAGAAAGTGGATGGGCTGGCTGCAAGGTACGCTCCGGCGTCAAACTCTCTTACCG 1755
QY 61 GAGGAGTCAATTCAGGTCGGGCTCAACCAATACCTGGTGGGTGCGACGTCCTCCATGC 120
DB 1756 GAGGAGTCAATTCAGGTCGGGCTCAACCAATACCTGGTGGGTGCGACGTCCTCCATGC 1815
QY 121 GAGCCCGAACCAGATAGCAAGTGTCACTTCATGCTGCAACGCCCTCCCATCACA 180

```

```

Db      1816 GAGCCGAAACCGGATGTAGACAGTGTCTCACTTCATCTCAGCGACCCCTCCACATCACA 1875
Qy      181  GAGAGACCGGCTAAGGCGAGGCTGTGGCGAGGGGCTCTCCCCCTCTCTGGCCAGCTCTTCA 240
Db      1876 GAGAGACCGGCTAAGGCGAGGCTGTGGCGAGGGGCTCTCCCCCTCTCTGGCCAGCTCTTCA 1935
Qy      241  GCTAGCCAGTGTCTGTGGCCCTCTCTGAAAGGCGACATACATTACCMAAATGACTTCCCA 300
Db      1936 GCTAGCCAGTGTCTGTGGCCCTCTCTGAAAGGCGACATACATTACCMAAATGACTTCCCA 1995
Qy      301  GACGCTGACCTCATCGAGGCCAACCTCTGTGGCGGATGAGATGGGCGGGGACATTAC 360
Db      1996 GACGCTGACCTCATCGAGGCCAACCTCTGTGGCGGATGAGATGGGCGGGGACATTAC 2055
Qy      361  GCGGTGAGTCAAGAAACAAGGTAGTAATCTCTGAGACTCTTTGACCCGCTCCGACGGAG 420
Db      2056 GCGGTGAGTCAAGAAACAAGGTAGTAATCTCTGAGACTCTTTGACCCGCTCCGACGGAG 2115
Qy      421  GAGATGAGCGGGAAGTGTCCGTCCCGGCGAGATCTCTGCGGAAATCCAAAGAAATTCCCA 480
Db      2116 GAGATGAGCGGGAAGTGTCCGTCCCGGCGAGATCTCTGCGGAAATCCAAAGAAATTCCCA 2175
Qy      481  CCAAGCATGCCCGCATGCGCACGCCCGGATTAACAACCTCCGCTGCTGGAATCTTGAAG 540
Db      2176 CCAAGCATGCCCGCATGCGCACGCCCGGATTAACAACCTCCGCTGCTGGAATCTTGAAG 2235
Qy      541  GCGCCGGAATAGTCCCTCAGTGTGATGATGAGTGGCCATCCCACTTAAGACCCCT 600
Db      2236 GCGCCGGAATAGTCCCTCAGTGTGATGATGAGTGGCCATCCCACTTAAGACCCCT 2295
Qy      601  CCTATACCACTTCAACGAGAAAGAGACAGTGTCTTACAGAAATCCACCGTCTTCT 660
Db      2296 CCTATACCACTTCAACGAGAAAGAGACAGTGTCTTACAGAAATCCACCGTCTTCT 2355
Qy      661  GCGCTGGCGAGCTTGCACAAAGCTTTGTGTAGCTCCGACCGTGGCGCTGACAGC 720
Db      2356 GCGCTGGCGAGCTTGCACAAAGCTTTGTGTAGCTCCGACCGTGGCGCTGACAGC 2415
Qy      721  GGCAGGGGAAACGCCCTCTCAACCAATCTCCAGACGAGGGGAGAGATCTACGTT 780
Db      2416 GGCAGGGGAAACGCCCTCTCAACCAATCTCCAGACGAGGGGAGAGATCTACGTT 2475
Qy      781  GAGTCGATTCCTCCATGCCCCCTTGAAGGGGAGCCCGGAGTCTCAGCGAC 840
Db      2476 GAGTCGATTCCTCCATGCCCCCTTGAAGGGGAGCCCGGAGTCTCAGCGAC 2535
Qy      841  GGGTCTGTGCTTACCGTGAAGAGAGCGCGGTGAGAGACGTGCTGTGCTGATGTC 900
Db      2536 GGGTCTGTGCTTACCGTGAAGAGAGCGCGGTGAGAGACGTGCTGTGCTGATGTC 2595
Qy      901  TACACATGAGACAGGGCTGTGATCAACGATCGCGTGGAGAAAGCAAGTGGCCATC 960
Db      2596 TACACATGAGACAGGGCTGTGATCAACGATCGCGTGGAGAAAGCAAGTGGCCATC 2655
Qy      961  AACGGCTTGAAGAACTTTGTGTGCTGACCAACAATGATCTAGCTTACCATCCGC 1020
Db      2656 AACGGCTTGAAGAACTTTGTGTGCTGACCAACAATGATCTAGCTTACCATCCGC 2715
Qy      1021 AGCGCAAGCCAGCGG 1035
Db      2716 AGCGCAAGCCAGCGG 2730

```

```

RESULT 6
AA046195
ID      AA046195 standard; cDNA to mRNA; 7065 BP.
AC      AA046195;
XX
XX      27-AUG-2003 (revised)
DT      25-MAR-2003 (revised)
DT      24-FEB-1994 (first entry)

```

```

XX      PT-NANBH virus non-structural proteins.
DE
XX      Parenterally transmitted non A non B hepatitis; PT-NANBH; NS4;
KM      hepatitis C virus; HCV; detection; diagnosis; antigen; vaccine;
KM      amplification; primer; polymerase chain reaction; PCR; ss.
XX      Hepatitis virus.
XX      OS
XX      WO9317110-A2.
XX      PD
XX      02-SEP-1993.
XX      PF
XX      19-FEB-1993; 93WO-GB000345.
XX      PR
XX      21-FEB-1992; 92GB-00003803.
XX      PA
XX      (WEIL ) WEILCOMBE FOUND LTD.
XX      PA
XX      Parker D, Rodgers BC;
XX      PI
XX      MPI; 1993-288415/36.
XX      DR
XX      P-ESDB; AAR41435.
XX      PT
XX      New recombinant polypeptide for diagnosing hepatitis C - contains three
XX      distinct antigens from different viral regions, also useful in protective
XX      vaccines.
XX      PS
XX      Example 1; Page 43-53; 99pp; English.
XX      CC
XX      The NS4 region from the 3' region of the PT-NANBH genome (AA046195) is
XX      amplified by PCR using primers D224 and D226 (AA046196-97) and the
XX      fragment (AA046198) is cloned into a vector and expressed in infected
XX      insect cells. The recombinant virus (BHC-13) was able to express the NS4
XX      specific recombinant protein at low levels in the infected insect cells.
XX      If at least three different PT-NANBH antigens are used to screen for PT-
XX      CC NANBH, the screening is much more sensitive as compared to the use of
XX      CC only two PT-NANBH antigens. Pref. antigens are described in AA046192-94.
XX      CC Two new antigenic regions of the PT-NANBH genome are given in AA046198-
XX      CC 99. AA046202 describes an improved PT-NANBH recombinant polypeptide.
XX      CC (Updated on 25-MAR-2003 to correct FN field.) (Updated on 27-AUG-2003 to
XX      CC correct OS field.)
XX      SQ      Sequence 7065 BP; 1468 A; 2123 C; 1994 G; 1480 T; 0 U; 0 other;

```

```

Query Match      99.1%; Score 1025.4; DB 2; Length 7065;
Best Local Similarity 99.4%; Pred. No. 1.4e-265;
Matches 1029; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

Qy      1  ACAGAGTGAATGGGGTGGCTGCGACAGTACGCTCCGCGTGCACAACTCTCCTACGG 60
Db      3919 ACAGAGTGAATGGGGTGGCTGCGACAGTACGCTCCGCGTGCACAACTCTCCTACGG 3978
Qy      61  GAGAGGTCACTTCAGTGTGGGCTCAACCAATACCTGTGTGGTGGTGGAGCTCCCATGC 120
Db      3979 GAGAGGTCACTTCAGTGTGGGCTCAACCAATACCTGTGTGGTGGTGGAGCTCCCATGC 4038
Qy      121  GAGCCCGAAACCGGATGTGAGAGTGTCTCACTTCATGTCTCAACGACCCCTCCCAATACA 180
Db      4039 GAGCCCGAAACCGGATGTGAGAGTGTCTCACTTCATGTCTCAACGACCCCTCCCAATACA 4098
Qy      181  GCAGAGACGGCTAAGCGCAGGCTGGCCAGGGGCTCTCCCCCTCTTGGCCAGCTCTTCA 240
Db      4099 GCAGAGACGGCTAAGCGCAGGCTGGCCAGGGGCTCTCCCCCTCTTGGCCAGCTCTTCA 4158
Qy      241  GCTAGCCAGTGTCTGTGGCCCTCTCTGAAAGGCGACATACATTACCMAAATGACTTCCCA 300
Db      4159 GCTAGCCAGTGTCTGTGGCCCTCTCTGAAAGGCGACATACATTACCMAAATGACTTCCCA 4218
Qy      301  GACGCTGACCTCATCGAGGCCAACCTCTGTGGCGGATGAGATGGGCGGGGACATTACC 360
Db      4219 GACGCTGACCTCATCGAGGCCAACCTCTGTGGCGGATGAGATGGGCGGGGACATTACC 4278

```


QY 361 CCGGTGAGTCAAGAAAGAGTAGTAATCTGGAAGCTTTTCGACCCGCTCCGAGCGAG 420
 DB 4279 CCGGTGAGTCAAGAAAGAGTAGTAATCTGGAAGCTTTTCGACCCGCTCCGAGCGAG 4338
 QY 421 GAGGATGACGGGAGAGTGTCCGTCCCGCGGAGATCTTCGGAAATTCAGAAATTTCCCA 480
 DB 4339 GAGGATGACGGGAGAGTGTCCGTCCCGCGGAGATCTTCGGAAATTCAGAAATTTCCCA 4398
 QY 481 CCAGCATGCCCCGATGGGACGCCCGGATTACACCTTCCTGCTGAGATCTTGGAG 540
 DB 4399 CCAGCATGCCCCGATGGGACGCCCGGATTACACCTTCCTGCTGAGATCTTGGAG 4458
 QY 541 GCCCGGAGTACGTCCCTCAGTGTATACATGGGATGCGACCTGACCTACTAAGACCCCT 600
 DB 4459 GCCCGGAGTACGTCCCTCAGTGTATACATGGGATGCGACCTGACCTACTAAGACCCCT 4518
 QY 601 CCTATACACCTTCACGAGAAAGAGACAGTTGTTCTGACAGAAATTCACCCGTCTTCT 660
 DB 4519 CCTATACACCTTCACGAGAAAGAGACAGTTGTTCTGACAGAAATTCACCCGTCTTCT 4578
 QY 661 GCTCTGCGGAGTGTCCGACAAAGGCTTTTGTAGTCTCCGAGACCTGCGCTGACAGC 720
 DB 4579 GCTCTGCGGAGTGTCCGACAAAGGCTTTTGTAGTCTCCGAGACCTGCGCTGACAGC 4638
 QY 721 GGCAGCGCAACCGCCCTCTGACCAATCTCCGACGAGCGGAGAGATCTGACGTT 780
 DB 4639 GGCAGCGCAACCGCCCTCTGACCAATCTCCGACGAGCGGAGAGATCTGACGTT 4698
 QY 781 GAGTGTATTCCTTCATATGCCCCCTTGAAGGGGAGACCGGAGACCCGATCTCAAGGAC 840
 DB 4699 GAGTGTATTCCTTCATATGCCCCCTTGAAGGGGAGACCGGAGACCCGATCTCAAGGAC 4758
 QY 841 GGGTCTTGTGTCTACCGTGAATGAGAGCGCGGTGAGAGAGCTGTCTGTCTGATGTCC 900
 DB 4759 GGGTCTTGTGTCTACCGTGAATGAGAGCGCGGTGAGAGAGCTGTCTGTCTGATGTCC 4818
 QY 901 TACACATGAGCAGAGCGCTGTATCAGCCCATGCGCTGCGGAGAAAGCAAGTCCCATC 960
 DB 4819 TACACATGAGCAGAGCGCTGTATCAGCCCATGCGCTGCGGAGAAAGCAAGTCCCATC 4878
 QY 961 AAGCGCTTGAACACTTTTGTGCGTCAACAAGATGTTCTACAGATCCCGC 1020
 DB 4879 AAGCGCTTGAACACTTTTGTGCGTCAACAAGATGTTCTACAGATCCCGC 4938
 QY 1021 AAGCGCAAGCCAGCGG 1035
 DB 4939 AAGCGCAAGCCAGCGG 4953
 RESULT 7
 ADD93733 standard; DNA; 7989 BP.
 ID ADD93733;
 AC ADD93733;
 XX 29-JAN-2004 (first entry)
 DT XX Hepatitis C virus strain J4 (BB7/J4NSB) replicon.
 DE XX Hepatitis C virus; virucide; ss.
 KM HCV; vaccine; virucide; ss.
 XX Hepatitis C virus.
 OS WO2003085084-A2.
 PN 16-OCT-2003.
 XX 03-APR-2003; 2003WO-US010177.
 PF 03-APR-2002; 2002US-036985P.
 PR 03-APR-2002; 2002US-036985P.
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 PA
 XX

PI Gates A, Gu B, Sarsky RT;
 XX WPI; 2003-804301/75.
 DR
 XX New hepatitis C virus (HCV) sub-genomic replicon, useful for facilitating
 PT screening or testing of anti-HCV drugs, comprises a nucleic acid
 PT construct encoding chimeric HCV non-structural proteins, and an NS5B
 PT polymerase gene.
 XX
 XX disclosure; Page 80-85; 159pp; English.
 PS
 XX
 XX The present sequence comprises a replicating hepatitis C virus (HCV) J4
 CC (BB7/J4NSB) replicon. The invention provides sub-genomic replicons of
 CC HCV comprising a nucleic acid construct encoding chimeric HCV
 CC nonstructural protein and an NS5B polymerase gene. A preferred replicon
 CC comprises an NS3 nucleotide sequence ADD93721 that encodes the first 75
 CC contiguous N-terminal amino acids of the NS3 of genotype 1b, of a BB7
 CC strain. A chimeric replicon may comprise an NS3 sequence from any of the
 CC 6 major HCV genotypes and subtypes but has its first 225 nucleotides of
 CC the coding sequence replaced by the BB7 strain NS3 sequence, especially
 CC where the replicon is from HCV genotype 1a (H77 strain) or genotype 1b
 CC (J4 strain). Stable cell lines expressing and replicating functional
 CC replicons containing sequences from HCV genotype 1a (strain H77) or
 CC genotype 1b (strain J4) within the prototype 1b replicon backbone from
 CC HCV strain BB7 are provided. These can be used to screen for compounds
 CC that modulate viral replication. The sub-genomic HCV replicon systems of
 CC the invention may provide the foundation for generating HCV replicons of
 CC all 6 major genotypes and subtypes to facilitate screening, testing and
 CC evaluating anti-infective agents for HCV disease(s).
 CC
 XX
 SQ Sequence 7989 BP; 1656 A; 2373 C; 2230 G; 1730 T; 0 U; 0 Other;
 Query Match 85.9%; Score 889.4; DB 10; Length 7989;
 Best Local Similarity 91.2%; Pred. No. 6.1e-229;
 Matches 944; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
 QY 1 ACAGAGTGGATGAGGAGTGGCGTGCACAGTACGCTCCGCGGCAACCTCTCTACGG 60
 DB 5092 ACAGAGTGGATGAGGAGTGGCGTGGTGCACAGTACGCTCCGCGGCAACCTCTCTACGG 5151
 QY 61 GAGGAGTCAATTCAGAGTGGCGTGCACAGTACGCTCCGCGGCAACCTCTCTACGG 120
 DB 5152 GAGGAGTCAATTCAGAGTGGCGTGGTGCACAGTACGCTCCGCGGCAACCTCTCTACGG 5211
 QY 121 GAGCCCGAACCAGATATGAGTGTCTCACTTCTCACTCTCAACGACCTCTCCATCA 180
 DB 5212 GAGCCCGAACCAGATATGAGTGTCTCACTTCTCACTCTCAACGACCTCTCCATCA 5271
 QY 181 GCAGAGAGGCTAAGGCGAGGCTGGCGAGGAGGCTCCCGCTCTTGGCAAGCTTTCA 240
 DB 5272 GCAGAGAGGCTAAGGCGAGGCTGGCGAGGAGGCTCCCGCTCTTGGCAAGCTTTCA 5331
 QY 241 GCTAGCAGTGTGTGGCCCTTCTTCAAGGAGCAATATGATCCCA 300
 DB 5332 GCTATCAGTGTGTGGCCCTTCTTCAAGGAGCAATATGATCCCA 5391
 QY 301 GACGCTGACCTATGAGGCGCAACTCTGTGTGGCGGAGATGAGTGGCGGAGATTA 360
 DB 5392 GACGCTGACCTATGAGGCGCAACTCTGTGTGGCGGAGATGAGTGGCGGAGATTA 5451
 QY 361 CGGCTGAGTCAAGAGCAAGGATGATCTGAGTCTTTGAGACCGCGCTCCGAGGAG 420
 DB 5452 CGGCTGAGTCAAGAGCAAGGATGATCTTGTGAGTCTTTGAGACCGCGCTCCGAGGAG 5511
 QY 421 GAGGATGACGGGAGAGTGTCCGTCCCGCGGAGATCTTCGGAAATTCAGAAATTTCCCA 480
 DB 5512 GAGGATGACGGGAGAGTGTCCGTCCCGCGGAGATCTTCGGAAATTCAGAAATTTCCCA 5571
 QY 481 CCAGCATGCCCCGATGGGACGCCCGGATTACACCTTCCTGCTGAGATCTTGGAG 540
 DB 5572 CCAGCATGCCCCGATGGGACGCCCGGATTACACCTTCCTGCTGAGATCTTGGAG 5631
 QY 541 GCCCGGAGTACGTCCCTCAGTGTATACATGGGATGCGACCTGACCTACTAAGACCCCT 600

```

Db      5632 GACCCGAGTACGTCCTCCAGGTGTACACGGGTGTCCATTGCCGCTCCAAAGGCCCT 5691
Qy      601 CCTATACCACTCCACGAGAAAGAGAGATGTTCTTCAAGAAATCCACCGTCTTCT 660
Db      5692 CCGATACCACTCCACGAGAAAGAGAGATGTTCTTCAAGAAATCCACCGTCTTCT 5751
Qy      661 GCCCTGCGAGAGCTTGCCAAAGAGCTTTGTAGTCCGAGACCGTCCGCTCCGACAGC 720
Db      5752 GCCTTGCGGAGGTGCGCAAAAGACCTTCGGAGCTCCGAATCGTCCGCTCCGACAGC 5811
Qy      721 GCGACGGCAACGCCCTCTCTGACCAATCTCTCCGACGAGCGAGAGAGATCTGACGTT 780
Db      5812 GCGACGGCAACGCCCTCTCTGACCAATCTCTCCGACGAGCGAGAGATCTGACGTT 5871
Qy      781 GAGGTGATTCCTCCATGCCCCCTTGAGGGGGAGCGGGGAGACCCCGATCTCAAGCGAC 840
Db      5872 GAGGTGATTCCTCCATGCCCCCTTGAGGGGGAGCGGGGAGATCTCCATCTCAAGCGAC 5931
Qy      841 GGGCTTGCTTACCGTGTAGTGAAGAGCGGGTGAAGAGAGTGTCTGTCTGATGTCC 900
Db      5932 GGGCTTGCTTACCGTGTAGTGAAGAGAGTGTCTGTCTGATGTCC 5991
Qy      901 TACACATGGAACAGCGGCTGTGATCACGCCATGCGCTGCGAGAGAAAGCAAGCTGCCATC 960
Db      5992 TACACATGGAACAGCGGCTGTGATCACGCCATGCGCTGCGAGAGAAAGTAAAGCTGCCATC 6051
Qy      961 AACGGGTGAGCACTTTGCTGCGTCAACCAACATGTTACGCTTACGATCCCGC 1020
Db      6052 AACCGGTGAGCACTTTGCTGCGTCAACCAACATGTTACGCTTACGATCCCGC 6111
Qy      1021 AGCGCAAGCCAGCGG 1035
Db      6112 AGCGCAAGCCCTCGG 6126

RESULT 8
ABK8574 ID ABK8574 standard; DNA; 8642 BP.
XX AC ABK8574;
XX DT 21-OCT-2002 (first entry)
XX DE Hepatitis C virus S 22-3 replicon.
XX KM Self-replicating; hepatitis C virus; HCV; HCV replication inhibitor;
XX OS Hepatitis C virus.
XX FH Key Location/Qualifiers
XX FT mutation /tag= a
XX FT CDS 1802..8407
XX FT /tag= b
XX FT /product= "HCV NS2-5B"
XX FT /note= "Viral enzymes"
XX FT mutation replace(4446,A)
XX FT /tag= c
XX FT mutation replace(5498,G)
XX FT /tag= d
XX FT mutation replace(6268,A)
XX FT /tag= e
XX PN WO200252015-A2.
XX PD 04-JUL-2002.
XX PF 20-DEC-2001; 2001WO-CA001843.
XX PR 22-DEC-2000; 2000US-0257857P.
XX
```

```

PA (BOEH ) BOEHRINGER INGELHEIM CANADA LTD.
XX Kukolj G, Pause A;
XX WPI: 2002-575382/61.
XX DR P-PSDB; ABG30582.
XX
PT New self-replicating RNA molecules from Hepatitis C virus (HCV), which
PT possess enhanced transduction or replication efficiency, useful for
PT evaluating potential inhibitors of HCV replication.
XX
PS Claim 11; Page 59-69; 140pp; English.
XX
CC The invention describes a self-replicating hepatitis C virus (HCV)
CC polynucleotide molecule comprising a 5'-non translated region (NTR),
CC where guanine at position 1 is substituted for adenine, a HCV polypeptide
CC region coding for a HCV polypeptide, and a 3'-NTR region. The self-
CC replicating Hepatitis C virus (HCV) RNA molecule is useful for evaluating
CC potential inhibitors of HCV replication. The HCV RNA molecule is also
CC useful for efficiently establishing cell culture replication. The self-
CC replicating polynucleotide molecule contains a 5'-NTR, where G at
CC position 1 is substituted for A, and therefore provides an alternative to
CC existing systems comprising a self-replicating HCV RNA molecule that, in
CC conjunction with mutations in the HCV non-structural region, such as the
CC G(2042)C/R mutations, transduces and/or replicates with greater
CC efficiency. This sequence represents hepatitis C virus replicon S22-3, a
CC self-replicating HCV polynucleotide molecule created from the replicon
CC APK12 (see ABK8573)
XX
SQ Sequence 8642 BP; 1768 A; 2557 C; 2438 G; 1877 T; 0 U; 2 Other;
```

```

Query Match 85.6%; Score 885.8; DB 6; Length 8642;
Best Local Similarity 90.9%; Pred. No. 5.9e-228;
Matches 941; Conservative 1; Mismatches 93; Indels 0; Gaps 0;
```

```

Qy      1 ACAGAGTGAATGGGCTGCGCTGCAAGTACGCTCCGCGGTGCAAACTCTCTTACGG 60
Db      5741 ACAGAGTGAATGGGCTGCGCTGCAAGTACGCTCCAGCGGTGCAAACTCTCTTACGG 5800
Qy      61 GAGGAGTCAATTCGAGGGGCTCAACCAATACCTGGTGGGTCGACGCTCCATGCG 120
Db      5801 GAGGAGTCAATTCGAGGGGCTCAATCAATACCTGGTGGGTCGACGCTCCATGCG 5860
Qy      121 GAGCCGGAACCGAGTGTAGTGTCTCACTTCAATGTCTACCGAGACCCCTCCACATCA 180
Db      5861 GAGCCGGAACCGAGTGTAGTGTCTCACTTCAATGTCTACCGAGACCCCTCCACATTA 5920
Qy      181 GCGAGACGCTTAACCGCAAGCTGCGCAGGGAGTCTCCCTCTCTTGGCCAGCTTTCA 240
Db      5921 GCGGAGACGCTTAACCGCTGAGCTGCGCAGGGAGTCTCCCTCTCTTGGCCAGCTTCA 5980
Qy      241 GCTAGCCAGTGTCTGCGCCCTTCTCTGGAAGCGCATATCAATTAACCAAAATGACTTCCA 300
Db      5981 GCTAGCCAGTGTCTGCGCCCTTCTCTGGAAGCGCATATCAATTAACCGTCAATGACTTCCA 6040
Qy      301 GACGCTGACCTCATCGAGGCCAACCTCTGTGTGCGCGCATGAGTGGCGGGAGATTAAC 360
Db      6041 GACGCTGACCTCATCGAGGCCAACCTCTGTGTGCGCGCATGAGTGGCGGGAGATTAAC 6100
Qy      361 CGCGTGAAGTCAAGAACAGAGTAAATCTTGAATCTTTTCAACCGCTCCGAGCGGAG 420
Db      6101 CGCGTGAAGTCAAGAAATTAAGTAAATTTTGGACTTTTCAAGCCGCTCCAGCGGAG 6160
Qy      421 GAGGATGAGCGGGAAGTCTCCGTCGCGGAGATCTCCGCGGAATCAAGAAATTTCCA 480
Db      6161 GAGGATGAGCGGGAAGTCTCCGTCGCGGAGATCTCCGCGAGTCAAGAAATTTCCCT 6220
Qy      481 CCAGGATGCGCGCATGCGGACGCCCGGATTAACAACCTCCGCTGCTGAGTCTCTGAGAG 540
Db      6221 CGAGGATGCGCGCATATGCGGACGCCCGGATTAACAACCTCCACTTTGAGTCTCTGAGAG 6280
Qy      541 GCCCGGACTAGTCTCTCAAGTGTATAGTGGTCCCACTGCACTTAAGACCCCT 600
```

Db 6281 GACCCGAGTACGTCCTCCAGTGTATACAGGGGATGTCATGCGCCCTGCCAGGCCCCCT 6340
Qy 601 CCTATACCACTTCCAGGAGAAAGAGACAGTTGTTGTACAGAAATCCACCGTGTCTTCT 660
Db 6341 CCGATACCACTTCCAGGAGAAAGAGACAGTTGTTGTGTCTGTACAAATCTTACCGTGTCTTCT 6400
Qy 661 GCGCTTGGAGGAGTTCGCAAAAGGCTTTGTAGTCCGGAGCCGTGCGGCGTGCAGC 720
Db 6401 GCGTTTGGAGGAGTTCGCAAAAGGCTTTGTAGTCCGGAGCCGTGCGGCGTGCAGC 6460
Qy 721 GCGAGCGGACCCGCTCTCTGACCAATCTCTCCAGCAGCGCGAGAGAGATCTGACGTT 780
Db 6461 GCGAGCGGACCCGCTCTCTCTGACCAATCTCTCCAGCAGCGCGAGAGATCTGACGTT 6520
Qy 781 GAGTGTGATTTCTTCATGCCCCCTTTGAGGGGAGCCGGGGAGACCCCGATCTTCAGGAC 840
Db 6521 GAGTGTGATTTCTTCATGCCCCCTTTGAGGGGAGCCGGGGAGATCCCGATCTTCAGGAC 6580
Qy 841 GGGTCTTGGTCTACCGGTGAGTGAAGAGCGCGGTGAGGAGCGTCTGCTGATGATGTC 900
Db 6581 GGGTCTTGGTCTACCGGTGAGTGAAGAGCGGTGAGGAGCGTCTGCTGATGATGTC 6640
Qy 901 TAGACATGACAGCGGCTCTGTATCAGCCATGCGCTGCGGAGGAGAAAGCAAGTCCCATC 960
Db 6641 TAGACATGACAGCGGCTCTGTATCAGCCATGCGCTGCGGAGGAGAAAGCAAGTCCCATC 6700
Qy 961 AAGCGCTTGAACACTCTTTTGTCTGCGTACACCAACATGCTTACCGTACCATCCGC 1020
Db 6701 AAGCGCTTGAACACTCTTTTGTCTGCGTACACCAACATGCTTACCGTACCATCCGC 6760
Qy 1021 AGCGCAAGCGGAGCGG 1035
Db 6761 AGCGCAAGCGGAGCGG 6775

RESULT 9

AAD25321 standard; cDNA; 7987 BP.

AAD25321;

12-MAR-2002 (first entry)

Hepatitis C virus (HCV) replBartman/delta2v/s cDNA.

Hepatitis C virus; HCV; transfection; infection; virus neutralisation;

gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;

Hepatitis C virus.

WO200189364-A2.

29-NOV-2001.

23-MAY-2001; 2001MO-US016822.

23-MAY-2000; 2000US-00576989.

(UNIV) UNIV WASHINGTON.

Rice CM, Blight KJ;

WPI; 2002-066755/09.

Hepatitis C virus variants having greater transfection efficiency and

ability to survive subpassage, useful as a vaccine for immunizing primate

to the virus, comprise non-naturally occurring viral sequences.

Claim 44; Page 66-69; 174pp; English.

The invention relates to Hepatitis C virus (HCV) variants which include

polynucleotides comprising non-naturally occurring HCV sequence and HCV

CC variants that have a transfection efficiency and ability to survive
CC subpassage greater than HCV that have wild-type polypeptide coding
CC regions. The polynucleotides of the invention are useful for identifying
CC a cell line that is permissive for infection with HCV and detecting
CC replication of HCV in cells of the cell line. They are also useful for
CC testing a compound for anti-viral properties and for inhibiting HCV
CC infection. They are also useful for the generation of defined HCV virus
CC stocks to develop in vitro and in vivo assays for virus neutralisation,
CC attachment, penetration and entry, structure/function studies on HCV
CC proteins and RNA elements and identification of new antiviral targets, a
CC systematic survey of cell culture systems and conditions to identify
CC those that support wild-type and variant HCV RNA replication and particle
CC release, production of adaptive HCV variants capable of more efficiency
CC replication in cell culture, production of HCV variants with altered
CC tissue or species tropism, establishment of alternative animal models for
CC inhibitor evaluation including those supporting HCV variant replication,
CC development of cell-free HCV replication assays, production of
CC immunogenic HCV particles for vaccination, engineering of attenuated HCV
CC derivatives as possible vaccine candidates, engineering of attenuated or
CC defective HCV derivatives for expression of heterologous gene products
CC for gene therapy and vaccine applications and for utilisation of the HCV
CC glycoproteins for targeted delivery of therapeutic agents to the liver
CC or other cell types with appropriate receptors. Vaccine comprising these
CC sequences is useful for inducing immunoprotection to HCV in a primate.
CC The present sequence is Hepatitis C virus (HCV) replBartman/delta2v/s
CC cDNA

SQ Sequence 7987 BP; 1647 A; 2368 C; 2243 G; 1729 T; 0 U; 0 Other;

Query Match 85.5%; Score 884.6; DB 6; Length 7987;

Best Local Similarity 90.9%; Pred. No. 1,2e-227;

Matches 941; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

Qy 1 ACAGAGTGTATGAGTGGGTGCGGCTGACAGGTACGCTCCGCGTGAACCTCTCTACCG 60
Db 5092 ACAGAGTGTATGAGTGGGTGCGGCTGACAGGTACGCTCCAGCGTGAACCTCTCTACCG 5151
Qy 61 GAGGAGTGTATGAGTGGGTGCGGCTGACAGGTACGCTCCGCGTGAACCTCTCTACCG 120
Db 5152 GAGGAGTGTATGAGTGGGTGCGGCTGACAGGTACGCTCCGCGTGAACCTCTCTACCG 5211
Qy 121 GAGCGCGAAGCGGATGTACAGTGTACATCTTCATCTGACGAGCCCTCCCATCACA 180
Db 5212 GAGCGCGAAGCGGATGTACAGTGTACATCTTCATCTGACGAGCCCTCCCATCACA 5271
Qy 181 GAGGAGTGTATGAGTGGGTGCGGCTGACAGGTACGCTCCGCGTGAACCTCTCTACCG 240
Db 5272 GAGGAGTGTATGAGTGGGTGCGGCTGACAGGTACGCTCCGCGTGAACCTCTCTACCG 5331
Qy 241 GCTAGGAGTGTATGAGTGGGTGCGGCTGACAGGTACGCTCCGCGTGAACCTCTCTACCG 300
Db 5332 GCTAGGAGTGTATGAGTGGGTGCGGCTGACAGGTACGCTCCGCGTGAACCTCTCTACCG 5391
Qy 301 GAGCGTGTATGAGTGGGTGCGGCTGACAGGTACGCTCCGCGTGAACCTCTCTACCG 360
Db 5392 GAGCGTGTATGAGTGGGTGCGGCTGACAGGTACGCTCCGCGTGAACCTCTCTACCG 5451
Qy 361 GCGGTGAGTGTATGAGTGGGTGCGGCTGACAGGTACGCTCCGCGTGAACCTCTCTACCG 420
Db 5452 GCGGTGAGTGTATGAGTGGGTGCGGCTGACAGGTACGCTCCGCGTGAACCTCTCTACCG 5511
Qy 421 GAGGATGAGCGGAGTGTATGAGTGGGTGCGGCTGACAGGTACGCTCCGCGTGAACCTCTCTACCG 480
Db 5512 GAGGATGAGCGGAGTGTATGAGTGGGTGCGGCTGACAGGTACGCTCCGCGTGAACCTCTCTACCG 5571
Qy 481 CAGGATGAGCGGAGTGTATGAGTGGGTGCGGCTGACAGGTACGCTCCGCGTGAACCTCTCTACCG 540
Db 5572 CAGGATGAGCGGAGTGTATGAGTGGGTGCGGCTGACAGGTACGCTCCGCGTGAACCTCTCTACCG 5631
Qy 541 GCGCGGAGTGTATGAGTGGGTGCGGCTGACAGGTACGCTCCGCGTGAACCTCTCTACCG 600
Db 5632 GCGCGGAGTGTATGAGTGGGTGCGGCTGACAGGTACGCTCCGCGTGAACCTCTCTACCG 5691

QY 601 CCTATACACCTTCCACGAGAAAAGAGACGTTGTTCTGACAGAAATCCACCGTGTCTTCT 660
CC (particularly preparation of attenuated HCV). The can also be used for
CC preparation of a liver-specific delivery system for gene therapy, and to
CC identify cells permissive for HCV replication. Virus RNA replicates
CC autonomously and with high efficiency in this cellular system, so that
CC variations in replication rates can be measured (for screening antiviral
CC agents) quantitatively or qualitatively, using standard laboratory
CC equipment. Efficient replication of HCV RNA is only achieved when the
CC specified RNA segments are present and when the transfected cells are
CC maintained under permanent selection pressure
XX
SQ Sequence 7989 BP; 1647 A; 2368 C; 2243 G; 1731 T; 0 U; 0 Other:
Query Match 85.5%; Score 884.6; DB 3; Length 7989;
Best Local Similarity 90.9%; Pred. No. 1.2e-227;
Matches 941; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
QY 1 ACAGAGTGAATGGGGTCCGCTGACAGTACGCTCCGGGCTGCAAACTCTCTACAGG 60
DB ACAGAGTGAATGGGGTCCGCTGACAGTACGCTCCGGGCTGCAAACTCTCTACAGG 5151
QY 61 GAGGAGTCAATTCAGATCGGGCTCAACCAATACCTGTTGGGTGACGTCCTACATG 120
DB GAGGAGTCAATTCAGATCGGGCTCAACCAATACCTGTTGGGTGACGTCCTACATG 5211
QY 5152 GAGGAGTCAATTCAGATCGGGCTCAACCAATACCTGTTGGGTGACGTCCTACATG 5211
DB GAGGAGTCAATTCAGATCGGGCTCAACCAATACCTGTTGGGTGACGTCCTACATG 5271
QY 121 GAGCCGGAACCGGATGAGAGTGTCTCACTTCACTGCTCAACGACCCCTCCACATCAG 180
DB GAGCCGGAACCGGATGAGAGTGTCTCACTTCACTGCTCAACGACCCCTCCACATCAG 5271
QY 5212 GAGCCGGAACCGGATGAGAGTGTCTCACTTCACTGCTCAACGACCCCTCCACATCAG 5271
DB GAGCCGGAACCGGATGAGAGTGTCTCACTTCACTGCTCAACGACCCCTCCACATCAG 5271
QY 181 GCAGAGAGCGCTAAGCGAGGCTGCGCAGGGGGTCTCCCTCTGTTGGCCAGCTTTCA 240
DB GCAGAGAGCGCTAAGCGAGGCTGCGCAGGGGGTCTCCCTCTGTTGGCCAGCTTTCA 5331
QY 5272 GCAGAGAGCGCTAAGCGAGGCTGCGCAGGGGGTCTCCCTCTGTTGGCCAGCTTTCA 5331
DB GCAGAGAGCGCTAAGCGAGGCTGCGCAGGGGGTCTCCCTCTGTTGGCCAGCTTTCA 5331
QY 241 GCTAGCGAGTGTCTGCGGCTTCTCCGAAAGGCAATACATTAACCAATGACTTCCA 300
DB GCTAGCGAGTGTCTGCGGCTTCTCCGAAAGGCAATACATTAACCAATGACTTCCA 5391
QY 5332 GCTAGCGAGTGTCTGCGGCTTCTCCGAAAGGCAATACATTAACCAATGACTTCCA 5391
DB GCTAGCGAGTGTCTGCGGCTTCTCCGAAAGGCAATACATTAACCAATGACTTCCA 5391
QY 301 GACGCTGACCTCATCGAGGCCAACCTCTGTGAGGAGTGAAGATGAGGAGGAGCAATTAC 360
DB GACGCTGACCTCATCGAGGCCAACCTCTGTGAGGAGTGAAGATGAGGAGGAGCAATTAC 5451
QY 5392 GACGCTGACCTCATCGAGGCCAACCTCTGTGAGGAGTGAAGATGAGGAGGAGCAATTAC 5451
DB GACGCTGACCTCATCGAGGCCAACCTCTGTGAGGAGTGAAGATGAGGAGGAGCAATTAC 5451
QY 361 GCGGTGAGTCAAGAGACAGAGTGAATCTCTGCACTTTTTCAGCCGCTCCAGCGGAG 420
DB GCGGTGAGTCAAGAGACAGAGTGAATCTCTGCACTTTTTCAGCCGCTCCAGCGGAG 5511
QY 5452 GCGGTGAGTCAAGAGACAGAGTGAATCTCTGCACTTTTTCAGCCGCTCCAGCGGAG 5511
DB GCGGTGAGTCAAGAGACAGAGTGAATCTCTGCACTTTTTCAGCCGCTCCAGCGGAG 5511
QY 421 GAGGATGAGCGGGAAGTGTCCGTCGCGGAGATCCGCGGAATCCAGAAATTCCTCA 480
DB GAGGATGAGCGGGAAGTGTCCGTCGCGGAGATCCGCGGAATCCAGAAATTCCTCA 5571
QY 5512 GAGGATGAGCGGGAAGTGTCCGTCGCGGAGATCCGCGGAATCCAGAAATTCCTCA 5571
DB GAGGATGAGCGGGAAGTGTCCGTCGCGGAGATCCGCGGAATCCAGAAATTCCTCA 5571
QY 481 CCAGCGATGCCCGCATGAGGACGCGCGGATTACACCTCTCGCTGAGTCTCGAAG 540
DB CCAGCGATGCCCGCATGAGGACGCGCGGATTACACCTCTCGCTGAGTCTCGAAG 5531
QY 5572 CGAGCGATGCCCGCATGAGGACGCGCGGATTACACCTCTCGCTGAGTCTCGAAG 5531
DB CGAGCGATGCCCGCATGAGGACGCGCGGATTACACCTCTCGCTGAGTCTCGAAG 5531
QY 541 GCCCGGACTACGTCCTCTCCAGTGTGATAGTGTGATGATGATGATGATGATGATGATGAT 600
DB GCCCGGACTACGTCCTCTCCAGTGTGATAGTGTGATGATGATGATGATGATGATGATGAT 5632
QY 5632 GACCCGGACTACGTCCTCTCCAGTGTGATAGTGTGATGATGATGATGATGATGATGATGAT 5632
DB GACCCGGACTACGTCCTCTCCAGTGTGATAGTGTGATGATGATGATGATGATGATGATGAT 5691
QY 601 CCTATACACCTTCCACGAGAAAAGAGACGTTGTTCTGACAGAAATCCACCGTGTCTTCT 660
DB CCTATACACCTTCCACGAGAAAAGAGACGTTGTTCTGACAGAAATCCACCGTGTCTTCT 5751
QY 5692 CGATATCACCTTCCACGAGAAAAGAGACGTTGTTCTGACAGAAATCCACCGTGTCTTCT 5751
DB CGATATCACCTTCCACGAGAAAAGAGACGTTGTTCTGACAGAAATCCACCGTGTCTTCT 5751
QY 661 GCCCTGGCGAGCTTCCCAAAAAGGTTTGTGAGTCCGGACCGTCCGCTCCGACAGC 720
DB GCCCTGGCGAGCTTCCCAAAAAGGTTTGTGAGTCCGGACCGTCCGCTCCGACAGC 5811
QY 5752 GCCCTGGCGAGCTTCCCAAAAAGGTTTGTGAGTCCGGACCGTCCGCTCCGACAGC 5811
DB GCCCTGGCGAGCTTCCCAAAAAGGTTTGTGAGTCCGGACCGTCCGCTCCGACAGC 5811
QY 721 GGCACGCGCAACGCGCCCTCTGACCAATCTCCGACGAGCGGAGCAAGATCTGACGTT 780
DB GGCACGCGCAACGCGCCCTCTGACCAATCTCCGACGAGCGGAGCAAGATCTGACGTT 5871
QY 5812 GGCACGCGCAACGCGCCCTCTGACCAATCTCCGACGAGCGGAGCAAGATCTGACGTT 5871
DB GGCACGCGCAACGCGCCCTCTGACCAATCTCCGACGAGCGGAGCAAGATCTGACGTT 5871
QY 781 GAGTGTATTCCTCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
DB GAGTGTATTCCTCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5931
QY 5872 GAGTGTATTCCTCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5931
DB GAGTGTATTCCTCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5931

CC agents for HCV infections, and to prepare vaccines against HCV infection
CC (particularly preparation of attenuated HCV). The can also be used for
CC preparation of a liver-specific delivery system for gene therapy, and to
CC identify cells permissive for HCV replication. Virus RNA replicates
CC autonomously and with high efficiency in this cellular system, so that
CC variations in replication rates can be measured (for screening antiviral
CC agents) quantitatively or qualitatively, using standard laboratory
CC equipment. Efficient replication of HCV RNA is only achieved when the
CC specified RNA segments are present and when the transfected cells are
CC maintained under permanent selection pressure
XX
SQ Sequence 7989 BP; 1647 A; 2368 C; 2243 G; 1731 T; 0 U; 0 Other:
Query Match 85.5%; Score 884.6; DB 3; Length 7989;
Best Local Similarity 90.9%; Pred. No. 1.2e-227;
Matches 941; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
QY 1 ACAGAGTGAATGGGGTCCGCTGACAGTACGCTCCGGGCTGCAAACTCTCTACAGG 60
DB ACAGAGTGAATGGGGTCCGCTGACAGTACGCTCCGGGCTGCAAACTCTCTACAGG 5151
QY 61 GAGGAGTCAATTCAGATCGGGCTCAACCAATACCTGTTGGGTGACGTCCTACATG 120
DB GAGGAGTCAATTCAGATCGGGCTCAACCAATACCTGTTGGGTGACGTCCTACATG 5211
QY 5152 GAGGAGTCAATTCAGATCGGGCTCAACCAATACCTGTTGGGTGACGTCCTACATG 5211
DB GAGGAGTCAATTCAGATCGGGCTCAACCAATACCTGTTGGGTGACGTCCTACATG 5271
QY 121 GAGCCGGAACCGGATGAGAGTGTCTCACTTCACTGCTCAACGACCCCTCCACATCAG 180
DB GAGCCGGAACCGGATGAGAGTGTCTCACTTCACTGCTCAACGACCCCTCCACATCAG 5271
QY 5212 GAGCCGGAACCGGATGAGAGTGTCTCACTTCACTGCTCAACGACCCCTCCACATCAG 5271
DB GAGCCGGAACCGGATGAGAGTGTCTCACTTCACTGCTCAACGACCCCTCCACATCAG 5271
QY 181 GCAGAGAGCGCTAAGCGAGGCTGCGCAGGGGGTCTCCCTCTGTTGGCCAGCTTTCA 240
DB GCAGAGAGCGCTAAGCGAGGCTGCGCAGGGGGTCTCCCTCTGTTGGCCAGCTTTCA 5331
QY 5272 GCAGAGAGCGCTAAGCGAGGCTGCGCAGGGGGTCTCCCTCTGTTGGCCAGCTTTCA 5331
DB GCAGAGAGCGCTAAGCGAGGCTGCGCAGGGGGTCTCCCTCTGTTGGCCAGCTTTCA 5331
QY 241 GCTAGCGAGTGTCTGCGGCTTCTCCGAAAGGCAATACATTAACCAATGACTTCCA 300
DB GCTAGCGAGTGTCTGCGGCTTCTCCGAAAGGCAATACATTAACCAATGACTTCCA 5391
QY 5332 GCTAGCGAGTGTCTGCGGCTTCTCCGAAAGGCAATACATTAACCAATGACTTCCA 5391
DB GCTAGCGAGTGTCTGCGGCTTCTCCGAAAGGCAATACATTAACCAATGACTTCCA 5391
QY 301 GACGCTGACCTCATCGAGGCCAACCTCTGTGAGGAGTGAAGATGAGGAGGAGCAATTAC 360
DB GACGCTGACCTCATCGAGGCCAACCTCTGTGAGGAGTGAAGATGAGGAGGAGCAATTAC 5451
QY 5392 GACGCTGACCTCATCGAGGCCAACCTCTGTGAGGAGTGAAGATGAGGAGGAGCAATTAC 5451
DB GACGCTGACCTCATCGAGGCCAACCTCTGTGAGGAGTGAAGATGAGGAGGAGCAATTAC 5451
QY 361 GCGGTGAGTCAAGAGACAGAGTGAATCTCTGCACTTTTTCAGCCGCTCCAGCGGAG 420
DB GCGGTGAGTCAAGAGACAGAGTGAATCTCTGCACTTTTTCAGCCGCTCCAGCGGAG 5511
QY 5452 GCGGTGAGTCAAGAGACAGAGTGAATCTCTGCACTTTTTCAGCCGCTCCAGCGGAG 5511
DB GCGGTGAGTCAAGAGACAGAGTGAATCTCTGCACTTTTTCAGCCGCTCCAGCGGAG 5511
QY 421 GAGGATGAGCGGGAAGTGTCCGTCGCGGAGATCCGCGGAATCCAGAAATTCCTCA 480
DB GAGGATGAGCGGGAAGTGTCCGTCGCGGAGATCCGCGGAATCCAGAAATTCCTCA 5571
QY 5512 GAGGATGAGCGGGAAGTGTCCGTCGCGGAGATCCGCGGAATCCAGAAATTCCTCA 5571
DB GAGGATGAGCGGGAAGTGTCCGTCGCGGAGATCCGCGGAATCCAGAAATTCCTCA 5571
QY 481 CCAGCGATGCCCGCATGAGGACGCGCGGATTACACCTCTCGCTGAGTCTCGAAG 540
DB CCAGCGATGCCCGCATGAGGACGCGCGGATTACACCTCTCGCTGAGTCTCGAAG 5531
QY 5572 CGAGCGATGCCCGCATGAGGACGCGCGGATTACACCTCTCGCTGAGTCTCGAAG 5531
DB CGAGCGATGCCCGCATGAGGACGCGCGGATTACACCTCTCGCTGAGTCTCGAAG 5531
QY 541 GCCCGGACTACGTCCTCTCCAGTGTGATAGTGTGATGATGATGATGATGATGATGATGAT 600
DB GCCCGGACTACGTCCTCTCCAGTGTGATAGTGTGATGATGATGATGATGATGATGATGATGAT 5632
QY 5632 GACCCGGACTACGTCCTCTCCAGTGTGATAGTGTGATGATGATGATGATGATGATGATGAT 5632
DB GACCCGGACTACGTCCTCTCCAGTGTGATAGTGTGATGATGATGATGATGATGATGATGATGAT 5691
QY 601 CCTATACACCTTCCACGAGAAAAGAGACGTTGTTCTGACAGAAATCCACCGTGTCTTCT 660
DB CCTATACACCTTCCACGAGAAAAGAGACGTTGTTCTGACAGAAATCCACCGTGTCTTCT 5751
QY 5692 CGATATCACCTTCCACGAGAAAAGAGACGTTGTTCTGACAGAAATCCACCGTGTCTTCT 5751
DB CGATATCACCTTCCACGAGAAAAGAGACGTTGTTCTGACAGAAATCCACCGTGTCTTCT 5751
QY 661 GCCCTGGCGAGCTTCCCAAAAAGGTTTGTGAGTCCGGACCGTCCGCTCCGACAGC 720
DB GCCCTGGCGAGCTTCCCAAAAAGGTTTGTGAGTCCGGACCGTCCGCTCCGACAGC 5811
QY 5752 GCCCTGGCGAGCTTCCCAAAAAGGTTTGTGAGTCCGGACCGTCCGCTCCGACAGC 5811
DB GCCCTGGCGAGCTTCCCAAAAAGGTTTGTGAGTCCGGACCGTCCGCTCCGACAGC 5811
QY 721 GGCACGCGCAACGCGCCCTCTGACCAATCTCCGACGAGCGGAGCAAGATCTGACGTT 780
DB GGCACGCGCAACGCGCCCTCTGACCAATCTCCGACGAGCGGAGCAAGATCTGACGTT 5871
QY 5812 GGCACGCGCAACGCGCCCTCTGACCAATCTCCGACGAGCGGAGCAAGATCTGACGTT 5871
DB GGCACGCGCAACGCGCCCTCTGACCAATCTCCGACGAGCGGAGCAAGATCTGACGTT 5871
QY 781 GAGTGTATTCCTCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
DB GAGTGTATTCCTCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5931
QY 5872 GAGTGTATTCCTCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5931
DB GAGTGTATTCCTCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5931

```
QY 841 GGGCTTGGTCTACCGGATGAGAGAGCGCGTGGAGAGCTCGTCTGCTGATGATCC 900
DB 5932 GGGCTTGGTCTACCGGATGAGAGAGCGCGTGGAGAGCTCGTCTGCTGATGATCC 9591
QY 901 TACACATGAGACAGCGCGCTGTGATACGCCATGCGGAGAGAAAGCAAGTCCCATC 960
DB 5992 TACACATGAGACAGCGCGCTGTGATACGCCATGCGGAGAGAAAGCAAGTCCCATC 6051
QY 961 AACGCGTTGAGCAACTTGTGCTGCTGATACCAACATGCTCTACGCTACCATCCCGC 1020
DB 6052 AATGCACTGAGCAACTTGTGCTGCTGATACCAACATGCTCTACGCTACCATCCCGC 6111
QY 1021 AGCGCAAGCAGCGCG 1035
DB 6112 AGCGCAAGCAGCGCG 6126

RESULT 11
AAD25322
ID AAD25322 standard; cDNA; 7989 BP.
XX
AC AAD25322;
XX
DT 12-MAR-2002 (first entry)
XX
DE Hepatitis C virus (HCV) replBartman/Avail cDNA.
XX
KM Hepatitis C virus; HCV; transfection; infection; virus neutralisation;
KW gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;
XX ss.
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT CDS 1801..7758
FT FT /*tag= a
FT FT /*product= "HCVreplBartman polypeptide"
FT FT misc_feature 7766
FT FT /*tag= b
FT FT /*note= "Nucleotide creating Avail site"
XX
PN WO200189364-A2.
XX
PD 29-NOV-2001.
XX
PF 23-MAY-2001; 2001WO-US016822.
XX
PR 23-MAY-2000; 2000US-00576989.
XX
PA (UNITV ) UNIV WASHINGTON.
XX
PI Rice CM, Blight KJ;
XX
DR MPI; 2002-066755/09.
XX
DR P-PSDB; AAB15717.
XX
PT Hepatitis C virus variants having greater transfection efficiency and
PT ability to survive subpassage, useful as a vaccine for immunizing primate
PT to the virus, comprise non-naturally occurring viral sequences.
XX
PS Claim 44; Page 69-71; 174pp; English.
XX
CC The invention relates to Hepatitis C virus (HCV) variants which include
CC polynucleotides comprising non-naturally occurring HCV sequence and HCV
CC variants that have a transfection efficiency and ability to survive
CC subpassage greater than HCV that have wild-type polypeptide coding
CC regions. The polynucleotides of the invention are useful for identifying
CC a cell line that is permissive for infection with HCV and detecting
CC replication of HCV in cells of the cell line. They are also useful for
CC testing a compound for anti-viral properties and for inhibiting HCV
CC infection. They are also useful for the generation of defined HCV virus
CC stocks to develop in vitro and in vivo assays for virus neutralisation,
CC attachment, penetration and entry, structure/function studies on HCV
```

```
CC protein and RNA elements and identification of new antiviral targets, a
CC systematic survey of cell culture systems and conditions to identify
CC those that support wild-type and variant HCV RNA replication and particle
CC release, production of adaptive HCV variants capable of more efficiency
CC replication in cell culture, production of HCV variants with altered
CC tissue or species tropism, establishment of alternative animal models for
CC inhibitor evaluation including those supporting HCV variant replication,
CC development of cell-free HCV replication assays, production of
CC immunogenic HCV particles for vaccination, engineering of attenuated HCV
CC derivatives as possible vaccine candidates, engineering of attenuated or
CC defective HCV derivatives for expression of heterologous gene products
CC for gene therapy and vaccine applications and for utilisation of the HCV
CC glycoproteins for targeted delivery of therapeutic agents to the liver
CC or other cell types with appropriate receptors. Vaccine comprising these
CC sequences is useful for inducing immunoprotection to HCV in a primate.
CC The present sequence is Hepatitis C virus (HCV) replBartman/Avail cDNA
XX
SQ Sequence 7989 BP; 1647 A; 2369 C; 2242 G; 1731 T; 0 U; 0 Other;
Query Match 85.5%; Score 884.6; DB 6; Length 7989;
Best Local Similarity 90.9%; Pred. No. 1.2e-227;
Matches 941; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
QY 1 ACAGAGTGGATGGGCTGCGCTGACAGAGTACGCTCCGCGCTGCAAACTCTCTACCG 60
DB 5092 ACAGAGTGGATGGGCTGCGCTGACAGAGTACGCTCCGCGCTGCAAACTCTCTACCG 5151
QY 61 GAGAGAGTCAATTCAGAGTGGCGGCTCAACCAATACCTGTTGGTGGAGCTCCCATC 120
DB 5152 GAGAGAGTCAATTCAGAGTGGCGGCTCAACCAATACCTGTTGGTGGAGCTCCCATC 5211
QY 121 GAGCCCGAACCAGATGTACAGTGTCTCACTTCAAGCTTCAACCGACCCCTCCCAATCA 180
DB 5212 GAGCCCGAACCAGATGTACAGTGTCTCACTTCAAGCTTCAACCGACCCCTCCCAATCA 5271
QY 181 GCAGAGACGGCTAAGGCGAGGCTGCGCAGAGGAGTCTCCCGCTCTTGGCCAGCTTCA 240
DB 5272 GCAGAGACGGCTAAGGCGAGGCTGCGCAGAGGAGTCTCCCGCTCTTGGCCAGCTTCA 5331
QY 241 GCTAGCAGTGTCTGCGCCCTTCTTGAAGGCGACATACATTACCAAAATGACTTCCCA 300
DB 5332 GCTAGCAGTGTCTGCGCCCTTCTTGAAGGCGACATACATTACCAAAATGACTTCCCA 5391
QY 301 GAGCTGACCTCATGAGGCGCAACCTCTCTGCGGCAATGAGGCGGGAATTAC 360
DB 5392 GAGCTGACCTCATGAGGCGCAACCTCTCTGCGGCAATGAGGCGGGAATTAC 5451
QY 361 CGGCTGAGTCAAGAGCAAGGTATGATCTGACTCTTTGACCCGCTCGAGCGGAG 420
DB 5452 CGGCTGAGTCAAGAGCAAGGTATGATCTGACTCTTTGACCCGCTCGAGCGGAG 5511
QY 421 GAGGATGAGCGGGAAGTGTCTGCGCGGAGATCTGCGGAAATCCAAAAATTTCCA 480
DB 5512 GAGGATGAGCGGGAAGTGTCTGCGCGGAGATCTGCGGAAATCCAAAAATTTCCA 5571
QY 481 CCAAGCATGCGCGCATGAGGCGCAAGGATTTCAACCTCTGCTGAGATCTTGGAG 540
DB 5572 CCAAGCATGCGCGCATGAGGCGCAAGGATTTCAACCTCTGCTGAGATCTTGGAG 5631
QY 541 GCCCGGACTACGTCCTCCTCAGTGTGATCACTGAGGCGCACTCACTAAGACCCCT 600
DB 5632 GCCCGGACTACGTCCTCCTCAGTGTGATCACTGAGGCGCACTCACTAAGACCCCT 5691
QY 601 CCTATACACCTTCAACGAGAAAGAGACAGTTGTTTGAAGAAATTCACGCTGCTTCT 660
DB 5692 CCTATACACCTTCAACGAGAAAGAGACAGTTGTTTGAAGAAATTCACGCTGCTTCT 5751
QY 661 GCCCTGGGAGACTTGCACAAAGGCTTTGTAGTCTCGACCGTGGCCCTTCCAGAC 720
DB 5752 GCCCTGGGAGACTTGCACAAAGGCTTTGTAGTCTCGACCGTGGCCCTTCCAGAC 5811
QY 721 GGCACGCGCAAGCGCCCTCTGACCAATCTCCAGAGACGCGGAGAGATCTGACGTT 780
```

Db 5812 GGCACGGACGCGCTCTCTCTGACGACGCTCCGACGACGCGGAGTCCGACGTT 5871
Qy 781 GAGTCGTATCTCTCCATGCGCCCTTTGAGGGGAGCCGGGAGACCCGATCTCAGCGAC 840
Db 5872 GAGTCGTATCTCTCCATGCGCCCTTTGAGGGGAGCCGGGAGATCCGATCTCAGCGAC 5931
Qy 841 GGGTCCTGCTCTACCGGTGAGTGAAGAGCGCGGTGAGACGTCGTCTGCTCGATGTC 900
Db 5932 GGGTCCTGCTCTACCGGTGAGTGAAGAGCGCGGTGAGACGTCGTCTGCTCGATGTC 5991
Qy 901 TACACATGACGACGCGCTCTGATCAGCGATGCGCTGCGAGAAAGCAAGTCCCATC 960
Db 5992 TACACATGACGACGCGCTCTGATCAGCGATGCGCTGCGAGAAAGCAAGTCCCATC 6051
Qy 961 AACGCGTTGAGCACTCTTGTGCGTCAACCAACATGCTACGCTACCATCCGCG 1020
Db 6052 AATGCACTGAGCACTCTTGTGCGTCAACCAACATGCTACGCTACCATCCGCG 6111
Qy 1021 AGCGCAAGCCGCGG 1035
Db 6112 AGCGCAAGCCGCGG 6126

RESULT 12

ADJ57845 standard; DNA; 7989 BP.

ADJ57845;

06-MAY-2004 (first entry)

HCV replicon encoding sequence.

hepatitis C virus; HCV; Antiinflammatory; Hepatotropic; Virucide; ds;
HCV replicon.

Unidentified.

Key Location/Qualifiers

FT CDS 1801..7759

/tag= a

/product= "HCV replicon"

12-AUG-2003; 2003WO-US025260.

12-AUG-2002; 2002US-0402661P.

(BRIM) BRISTOL-MYERS SQUIBB CO.

Gao M, Lemm JA, O'Boyle DR, Nower P, Rigat K, Sun J;

WPI; 2004-180685/17.

P-PSDB; ADJ57846.

Use of hepatitis C virus assays or reporter assays, e.g. identifying a
compound that inhibits hepatitis C virus RNA replication or identifying a
compound that modulates the activity of a gene of interest.

Claim 3; SEQ ID NO 1; 45pp; English.

The present invention relates to the use of hepatitis C virus (HCV)
assays for identifying a compound that inhibits HCV RNA replication and
reporter assays for identifying a compound that modulates the activity of
a gene of interest. The assays are useful for identifying a compound that
inhibits HCV RNA replication or for identifying a compound that modulates
the activity of a gene of interest. The HCV assay is useful for high
throughput screening that quantifies both the amount of HCV RNA
replication inhibitory activity associated with a test compound and the
amount of cytotoxicity associated with the test compound. The compound is

CC useful for treating hepatitis C infection. Assays of the invention have
CC distinct advantages when compared to qRT-PCR or other methods in that
CC assays of the invention may take place in situ in a detector based crude
CC cell lysate, which requires no further preparation prior to performing
CC the assays. The assays do not also involve numerous manipulations to add
CC or subtract reagents after addition of test compounds and are desirably
CC based on a viral protein which is required by the HCV replicon for
CC replication. The present sequence represents a HCV replicon encoding
CC sequence used in the assay of the invention.

SQ Sequence 7989 BP; 1647 A; 2368 C; 2243 G; 1731 T; 0 U; 0 Other;

Query Match 85.5%; Score 884.6; DB 12; Length 7989;

Best Local Similarity 90.9%; Pred. No. 1.2e-227;

Matches 941; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

Qy 1 ACAAGAGTGAATGGGGTGGGCTGACAGGTAACGTCGGGGTGCMAACCTCTCTTACGG 60
Db 5092 ACAGAAAGTGAATGGGGTGGGCTGACAGGTAACGTCGGGGTGCMAACCTCTCTTACGG 5151
Qy 61 GAGGAGTCAATTCAGGTCGAGGCTCAACCAATACCTGTTGGGTGCGACGTCATGCG 120
Db 5152 GAGGAGTCAATTCAGGTCGAGGCTCAACCAATACCTGTTGGGTGCGACGTCATGCG 5211
Qy 121 GAGCCCGAACCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 180
Db 5212 GAGCCCGAACCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 5271
Qy 181 GCAAGAGAGGCTTAAGCGGAGGCTGCGCAGGGGGTCTCCCGCTCTTGGCCAGCTTTCA 240
Db 5272 GCGGAGAGGCTTAAGCGGAGGCTGCGCAGGGGGTCTCCCGCTCTTGGCCAGCTTTCA 5331
Qy 241 GCTAGCCAGTGTGTCTGGGCTTCTCCGAAAGCGCATATATACCAAAATGACTTTCCA 300
Db 5332 GCTAGCCAGTGTGTCTGGGCTTCTCCGAAAGCGCATATATACCAAAATGACTTTCCA 5391
Qy 301 GACGCTGACCTCATGAGAGGCAACCTCTGTGGCGGATGAAGATGGGCGGAGCATTAAC 360
Db 5392 GACGCTGACCTCATGAGAGGCAACCTCTGTGGCGGATGAAGATGGGCGGAGCATTAAC 5451
Qy 361 GCGGTGAGTCAAGAGAAAGATAGTAACTCTGGAATCTTTTCCGACCCGCTCCGAGCGAG 420
Db 5452 GCGGTGAGTCAAGAGAAAGATAGTAACTCTTTTCCGAGCCGCTCCGAGCGAG 5511
Qy 421 GAGGATGACGAGAAAGTGTCCGTCGCGGAGATCTCGCGAAATCCAGAAATTTCCA 480
Db 5512 GAGGATGACGAGAAAGTGTCCGTCGCGGAGATCTCGCGAAATTTCCA 5571
Qy 481 CCAGGATGCCCGCATGGGCAAGCCCGGATTACAACTCTCGCTGAGTCTCTGAAG 540
Db 5572 CGAGGATGCCCGCATGGGCAAGCCCGGATTACAACTCTCGCTGAGTCTCTGAAG 5631
Qy 541 GCCCGGATCAAGTCTCTCCAGTGTATAGGTGCGCACTGCCAATCTTAAGACCCCT 600
Db 5632 GACCGGATCAAGTCTCTCCAGTGTATAGGTGCGCACTGCCAATCTTAAGACCCCT 5691
Qy 601 CCTATACCACTCAACGAGAAAGAGACAGTGTCTGACAGATCCACGCTGCTCT 660
Db 5692 CGATACCACTCAACGAGAAAGAGACAGTGTCTGACAGATCCACGCTGCTCT 5751
Qy 661 GCCCTGGGAGCTTGCACAAAGCTTTTGTAGCTCCGAGACGTCGCGCGTGCAGACG 720
Db 5752 GCTTGGGAGCTTGCACAAAGCTTTTGTAGCTCCGAGACGTCGCGCGTGCAGACG 5811
Qy 721 GGCACGGACACCGCCCTCTCTGACCAATCTCTCCAGACGCGGAGAGATCTGACGTT 780
Db 5812 GGCACGGACACCGCCCTCTCTGACCAATCTCTCCAGACGCGGAGAGATCTGACGTT 5871
Qy 781 GAGTCGTATCTCTCCATGCGCCCTTTGAGGGGAGCCGGGAGATCCGATCTCAGCGAC 840
Db 5872 GAGTCGTATCTCTCCATGCGCCCTTTGAGGGGAGCCGGGAGATCCGATCTCAGCGAC 5931
Qy 841 GGGTCCTGCTCTACCGGTGAGTGAAGAGCGCGGTGAGACGTCGTCTGCTCGATGTC 900

|||||
Db 5932 GGGTCTGGTCTACCGGTAACGAGAGGCTAGTAGAGACCTCGTCTGCTGATGTCC 5991
Qy 901 TACACATGGAACAGCGCTGTGATCAAGCCATGCGCTGCGAGAAAGCAAGTCCCATC 960
Db 5992 TACACATGGAACAGCGCTGTGATCAAGCCATGCGCTGCGAGAAAGCAAGTCCCATC 6051
Qy 961 AACCGTTGAGCAACTTTGTGCTGCGTCAACCAACATGCTTACGCTACACATCCCGC 1020
Db 6052 AATGCACTGAGCAACTTTGTGCTGCGTCAACCAACATGCTTATGTATCAACATCTGCG 6111
Qy 1021 AGCGCAAGCCAGCGG 1035
Db 6112 AGCGCAAGCTGCGG 6126

RESULT 13
AAL47279
ID AAL47279 standard; DNA; 7991 BP.
XX AAL47279;
XX 30-AUG-2002 (first entry)
XX Hepatitis C virus sub-genomic replicon recombinant clone HCV9.
XX Hepatitis C virus core-neo; NS3 proteinase/helicase; vaccine; diagnosis;
XX virucide; hepatotropic; gene therapy; anti-viral; gene; ds.
XX Hepatitis C virus.
XX WO200238793A2.
XX 16-MAY-2002.
XX 02-NOV-2001; 2001WO-US046350.
XX 07-NOV-2000; 2000US-0245866P.
XX
XX (ANAD-) ANADYS PHARM INC.
XX Bichko V;
XX
XX MPI; 2002-49082/52.
XX
XX Novel nucleic acid encoding replication competent recombinant hepatitis C
XX virus genome useful for screening anti-hepatitis C virus therapeutics and
XX for vaccine development.
XX
XX Claim 9; Page 61-65; 85pp; English.
XX
XX The present invention provides protein and coding sequences from
XX Hepatitis C virus (HCV), comprising all or part of the HCV genome and
XX able to replicate efficiently when transfected into a susceptible cell
XX line without reducing the growth rate of the cell line by more than 10
XX fold. The sequences are useful for screening for anti-HCV therapeutics,
XX for detecting antibodies to HCV in a biological sample such as blood,
XX serum, plasma, blood cells, lymphocytes, or liver cells from a subject,
XX for deriving authentic HCV components such as replication-competent non-
XX infectious, replication-defective infection-component, and replication-
XX defective non-infectious HCV, in gene therapy or gene vaccination
XX targeted to hepatic tissue for treating an animal infected or susceptible
XX to HCV infection and for studying HCV infection and propagation. The
XX present sequence is a clone of a fragment of the HCV genome designated
XX HCV93
XX
XX Sequence 7991 BP; 1647 A; 2369 C; 2245 G; 1730 T; 0 U; 0 Other;
XX
XX Query Match 85.5%; Score 884.6; DB 6; Length 7991;
XX Best Local Similarity 90.9%; Pred. No. 1.2e-227;
XX Matches 941; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
Qy 1 ACAGAGTGATGGGTGGCTGACAGGTACGCTCCGCGTGCAAACTCTCTACGG 60

|||||
Db 5091 ACAGAGTGATGGGTGGCTGACAGGTACGCTCCGCGTGCAAACTCTCTACGG 5150
Qy 61 GAGAGGTCACTTCCAGGTGCGGCTCAACCAATACCTGTTGGGTGCGAGCTCCCATC 120
Db 5151 GAGAGGTCACTTCCAGGTGCGGCTCAACCAATACCTGTTGGGTGCGAGCTCCCATC 5210
Qy 121 GAGCCGAAACCGGATGATGAGAGTCACTTCCATGCTCAACCGACCCCTCCCATCA 180
Db 5211 GAGCCGAAACCGGATGATGAGAGTCACTTCCATGCTCAACCGACCCCTCCCATCA 5270
Qy 181 GCAGAGCGGCTTAAGCGCAAGCTGCGCAGAGGGGTCTCCCTCTTGGCCAGCTTTCA 240
Db 5271 GCGGAGCGGCTTAAGCGCTAGGCTAGGCTGCGCAGAGGGATCTCCCTCTTGGCCAGCTATCA 5330
Qy 241 GCTAGCCAGTTGTCTGCGCTTCTCTGAAAGCGCATTAATCCAAATGACTTCCCA 300
Db 5331 GCTAGCCAGCTGTCTGCGCTTCTCTGAAAGCGCATTAATCCGTCATGACTCCCG 5390
Qy 301 GACGCTGACCTCATGAGGCGCAACCTCTGTCGCGCATGAGATGGCGGAGCATTAAC 360
Db 5391 GACGCTGACCTCATGAGGCGCAACCTCTGTCGCGCATGAGATGGCGGAGCATTAAC 5450
Qy 361 CGCGTGAAGTCAAGAAACAGGTAGTAATCTGAACTTTTGGACCCGCTCGAGCGAG 420
Db 5451 CGCGTGAAGTCAAGAAATAGGTAGTAATTTTGGACTTTTGGAGCGCTCCAGCGAG 5510
Qy 421 GAGGATGAGCGGGAAGTGTCCGTCGCGCGAGATCTTCGGAATTCAGAAATTC 480
Db 5511 GAGGATGAGCGGGAAGTATCCGTTCCGCGGAGATCTTCGGAAGTTCAGAAATTC 5570
Qy 481 CCAGCGATGCCGATGAGGCGACCGCGGATTCACCCCTCGCTGAGTCTGGAAG 540
Db 5571 CGAGCGATGCCGATGAGGCGACCGCGGATTCACCCCTCACTGTAAGTCTGGAAG 5630
Qy 541 GCCCGGATTAAGTCTCCCTCAAGTGTATCAAGGTCGCCACTGCACTTAAGACCCCT 600
Db 5631 GACCGGATTAAGTCTCCCTCAAGTGTATCAAGGTCGCCACTGCACTTAAGACCCCT 5690
Qy 601 CCTATACCACTTCCAGCGAAGAAAGAGACAGTTGTTCTGACAGAAATCCACCGTCTTCT 660
Db 5691 CCGATACCACTTCCAGCGAAGAAAGAGACAGTTGTTCTGACAGAAATCCACCGTCTTCT 5750
Qy 661 GCCCTGCGGAGCTTGTCCCAAAAGGCTTTGTAGTCCGACCGTCCGCGCTCCAGCAG 720
Db 5751 GCTTGGCGGAGCTGCGCAAAAGACCTTCCGAGCTCCGAATCTGCGCGCTCCAGCAG 5810
Qy 721 GGCACGCGCAACCGCCCTCTCTGACCAAGCTTCCGACGAGCGGAGCAGATCTGAGTT 780
Db 5811 GGCACGCGCAACCGCCCTCTCTGACCAAGCTTCCGACGAGCGGAGATCTGAGTT 5870
Qy 781 GAGTGTATCTCTCATGTCGCCCTCTGAGGGGAGCGCGGGAGACCCGATCTCAGCGAC 840
Db 5871 GAGTGTATCTCTCATGTCGCCCTCTGAGGGGAGCGCGGGAGATCTCAGCGAC 5930
Qy 841 GGGTCTTGTCTTACCGTGAAGTGAAGGCGGCTGAGAGCTGTCTGCTCTGATGTCC 900
Db 5931 GGGTCTTGTCTTACCGTGAAGTGAAGGCGGCTGAGAGCTGTCTGCTCTGATGTCC 5990
Qy 901 TACACATGGAACAGCGCTGTGATCAAGCCATGCGCTGCGAGAAAGCAAGTCCCATC 960
Db 5991 TACACATGGAACAGCGCTGTGATCAAGCCATGCGCTGCGAGAAAGCAAGTCCCATC 6050
Qy 961 AACCGTTGAGCAACTTTGTGCTGCGTCAACCAACATGCTTACGCTACACATCCCGC 1020
Db 6051 AATGCACTGAGCAACTTTGTGCTGCGTCAACCAACATGCTTATGTATCAACATCTGCG 6110
Qy 1021 AGCGCAAGCCAGCGG 1035
Db 6111 AGCGCAAGCTGCGG 6126

RESULT 14


```
AA147276
ID AA147276 standard; DNA; 7992 BP.
AC AA147276;
DT 30-AUG-2002 (first entry)
XX
DE Hepatitis C virus sub-genomic replicon clone I377-NS3-3'UTR.
XX
KM Hepatitis; HCV, core-neo, NS3 proteinase/helicase; vaccine; diagnosis;
KM virucide; hepatotropic; gene therapy; anti-viral; gene; ds.
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT CDS 342..1181
FT /tag= a
FT /product= "core-neo fusion protein"
FT CDS 1801..7758
FT /tag= b
FT /product= "NS3 proteinase/helicase"
XX
PN WO200238793-A2.
XX
PD 16-MAY-2002.
XX
PF 02-NOV-2001; 2001WO-US046350.
XX
PR 07-NOV-2000; 2000US-0245866P.
XX
PA (ANAD-) ANADYS PHARM INC.
XX
PI Bichko V;
XX
DR WPI; 2002-490082/52.
DR P-PSDB; AAO18000, AAO18001.
XX
PT Novel nucleic acid encoding replication competent recombinant hepatitis C
PT virus genome useful for screening anti-hepatitis C virus therapeutics and
PT for vaccine development.
XX
PS Claim 6; Page 43-47; 85pp; English.
XX
CC The present invention provides protein and coding sequences from
CC Hepatitis C virus (HCV), comprising all or part of the HCV genome and
CC able to replicate efficiently when transfected into a susceptible cell
CC line without reducing the growth rate of the cell line by more than 10
CC fold. The sequences are useful for screening for anti-HCV therapeutics,
CC for detecting antibodies to HCV in a biological sample such as blood,
CC serum, plasma, blood cells, lymphocytes, or liver cells from a subject,
CC for deriving authentic HCV components such as replication-competent non-
CC infectious, replication-defective infection-component, and replication-
CC defective non-infectious HCV, in gene therapy or gene vaccination
CC targeted to hepatic tissue for treating an animal infected or susceptible
CC to HCV infection and for studying HCV infection and propagation. The
CC present sequence is a clone of a fragment of the HCV genome which encodes
CC the core-neo and NS3 proteinase/helicase proteins
XX
SQ Sequence 7992 BP; 1648 A; 2369 C; 2243 G; 1732 T; 0 U; 0 Other;
Query Match 85.5%; Score 884.6; DB 6; Length 7992;
Best Local Similarity 90.9%; Pred. No. 1.2e-227;
Matches 941; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
QY 1 ACAGAGTGGATGGGTCGGCTGCAAGTACGCTCCGCGTGAACCTCTCTTACCG 60
DB 5092 ACAGAGTGGATGGGTCGGCTGCAAGTACGCTCCGCGTGAACCTCTCTTACCG 5151
QY 61 GAGGAGTGCATCTTCCAGTTCGGGCTCAACCAATATCTGTTGGTTCGACGCTCCATCC 120
DB 5152 GAGGAGTGCATCTTCCAGTTCGGGCTCAACCAATATCTGTTGGTTCGACGCTCCATCC 5211
QY 121 GAGCCGGAACCGGATGAGAGTGTCACTTCATGCTCAACGACCCCTCCACATCA 180
```

```
DB 5212 GAGCCGGAACCGGATGAGAGTGTCTCACTTCATGCTCAACGACCCCTCCACATTAACG 5271
QY 181 GCAAGAGAGGCTTAAGCCGAGGCTGCGCCAGGGGGTCTCCCCCTCTTGGCCAGCTCTTCA 240
DB 5272 GCGGAGAGGCTTAAGCCGAGGCTGCGCCAGGGGGTCTCCCCCTCTTGGCCAGCTCTTCA 5331
QY 241 GCTAGCCAGTGTCTGCGCCCTTCTCGAAGGCGCATATCATTAACCAATATGACTTCCCA 300
DB 5332 GCTAGCCAGTGTCTGCGCCCTTCTCGAAGGCGCATATCATTAACCAATATGACTTCCCG 5391
QY 301 GACGCTGACCTTATGAGAGCCCAACCTCTGTGGCGGATGAGATGGCGGGACATTAAC 360
DB 5392 GACGCTGACCTTATGAGAGCCCAACCTCTGTGGCGGATGAGATGGCGGGACATTAAC 5451
QY 361 CCGCTGAGTCAAGAGAAACAAGTATTAATCTGATCTTTTCGACCCGCTCCGAGCGGAG 420
DB 5452 CCGCTGAGTCAAGAGAAACAAGTATTAATCTTTCGATCTTTTCGAGCCGCTCAAGCGGAG 5511
QY 421 GAGGATGAGCGGGAAGTGTCCGTCGCCGCGGAGATCTCGCGGAATCCAGAAATTTCCCA 480
DB 5512 GAGGATGAGAGGGAAGTATCCGTTCCGCGGAGATCTCGCGGAGTCCAGGAATTTCCCT 5571
QY 481 CCAGGATGCCCGCATYGGGCAAGCCCGGATTACAACCTCTCGCTGAGTCTCTGAAAG 540
DB 5572 CGAGGATGCCCGCATYATGGGCAAGCCCGGATTACAACCTCTCACTGTAGTCTCTGAAAG 5631
QY 541 GCCCGGATCTAGTCTCTCTCCAGTGTATATAGGTGCCCACTCCATCTTAAGACCCCT 600
DB 5632 GACCCGGACTAGTCTCTCTCCAGTGTATATAGGTGCCCACTCTTAAGACCCCT 5691
QY 601 CCTATACCACTTCAACGAGAAAGAGAGCAGTTGTCTGACAGATCCACCGTGTCTTCT 660
DB 5692 CGATATCACTTCAACGAGAAAGAGAGCAGTTGTCTGATCAAGATCTAACCTGTCTTCT 5751
QY 661 GCCCTGCGGAGCTTGGCCACAAGAGCTTTGTAGCTTCCGACCGTGGCCGTGCACAGC 720
DB 5752 GCTTGGGAGAGCTGCGCAAAAGACCTTGGGAGCTCGAATGTCGCGCGTGCACAGC 5811
QY 721 GGCACGGAACCGCCCTCTCTGACCAATCTTCCAGCAGCGCGAGAGATCTGACGTT 780
DB 5812 GGCACGGAACCGCCCTCTCTGACCAATCTTCCAGCAGCGCGAGATCTGACGTT 5871
QY 781 GAGTGTATTTCTTCATGACCCCTTTGAGGGGAGCGGGGAGACCCGATCTGACGAC 840
DB 5872 GAGTGTATTTCTTCATGACCCCTTTGAGGGGAGCGGGGAGATCCGATCTGACGAC 5931
QY 841 GGGTCTTGTCTTACCGTGAAGTGAAGAGCCGATGAGACGTCTGCTGCTGATGTC 900
DB 5932 GGGTCTTGTCTTACCGTGAAGTGAAGAGCCGATGAGACGTCTGCTGCTGATGTC 5991
QY 901 TACATATGACAGCGGCTCTGATCAAGCATGCGCTGCGAGAGAAAGAGAGTCCCATC 960
DB 5992 TACATATGACAGCGGCTCTGATCAAGCATGCGCTGCGAGAGAAAGAGAGTCCCATC 6051
QY 961 AACGGTTGAGCAACTTTTGTGTGTCACCAACAATGCTAGCTACCAATCCGCG 1020
DB 6052 AATGCACTGAGCAACTTTTGTGTGTCACCAACAATGCTAGCTACCAATCTTGC 6111
QY 1021 AGCGCAACCGAGCGG 1035
DB 6112 AGCGCAACCGAGCGG 6126
RESULT 15
AA147280
ID AA147280 standard; DNA; 7992 BP.
XX
AC AA147280;
DT 30-AUG-2002 (first entry)
XX
KM Hepatitis C virus sub-genomic replicon recombinant clone HCVR22.
XX
```


XX Hepatitis; HCV; core-neo; NS3 proteinase/helicase; vaccine; diagnosis;
 KM virucide; hepatotropic; gene therapy; anti-viral; gene; ds.
 XX Hepatitis C virus.
 OS WO200238793-A2.
 XX
 PN 16-MAY-2002.
 XX
 PD 02-NOV-2001; 2001WO-US046350.
 XX
 PF 07-NOV-2000; 2000US-0245866P.
 XX
 PR (AMND-) AMNDYS PHARM INC.
 XX
 PA Bichko V;
 XX
 PI WPI; 2002-490082/52.
 XX
 DR Novel nucleic acid encoding replication competent recombinant hepatitis C
 PT virus genome useful for screening anti-hepatitis C virus therapeutics and
 PT for vaccine development.
 XX
 PS Claim 10; Page 66-70; 85pp; English.
 XX
 CC The present invention provides protein and coding sequences from
 CC Hepatitis C virus (HCV), comprising all or part of the HCV genome and
 CC able to replicate efficiently when transfected into a susceptible cell
 CC line without reducing the growth rate of the cell line by more than 10
 CC fold. The sequences are useful for screening for anti-HCV therapeutics,
 CC for detecting antibodies to HCV in a biological sample such as blood,
 CC serum, plasma, blood cells, lymphocytes, or liver cells from a subject,
 CC for deriving authentic HCV components such as replication-competent non-
 CC infectious; replication-defective infection-component, and replication-
 CC defective non-infectious HCV, in gene therapy or gene vaccination
 CC targeted to hepatic tissue for treating an animal infected or susceptible
 CC to HCV infection and for studying HCV infection and propagation. The
 CC present sequence is a clone of a fragment of the HCV genome designated
 CC HCVR22
 CC
 SQ Sequence 7992 BP; 1646 A; 2368 C; 2245 G; 1733 T; 0 U; 0 Other;

Query Match 85.5%; Score 884.6; DB 6; Length 7992;
 Best Local Similarity 90.9%; Pred. No. 1.2e-227;
 Matches 941; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 1 ACAGAGTGGATGGGTGGCGGTGACAGGTACGCTCCGCGGTGCAAACTCTCTACCG 60
 DB 5092 ACGAAGTGGATGGGTGGCGGTGACAGGTACGCTCCGCGGTGCAAACTCTCTACCG 5151
 QY 61 GAGGAGTGCATTCAGGTGCGGCTCAACCAATACCTGTTGGTGCAGCTCCATGC 120
 DB 5152 GAGGAGTGCATTCAGGTGCGGCTCAATCAATACCTGTTGGTGCAGCTCCATGC 5211
 QY 121 GAGCCCGAACCAGATATGACAGTCTCACTTCAATGCTACCGACCCCTCCACATACA 180
 DB 5212 GAGCCCGAACCAGATATGACAGTCTCACTTCAATGCTACCGACCCCTCCACATACA 5271
 QY 181 GCAGAGACGGCTAAGCGCAGAGCTGCGCAGAGGGGTCTCCCTCTCTGGCAGACTCTTCA 240
 DB 5272 GCGGAGACGGCTAAGCGTACGCTGCGCAGAGGGGTCTCCCTCTCTGGCAGACTCTTCA 5331
 QY 241 GCTAGCCAGTTGTCTGCGCTTCTCTGAAAGCGACATTAATCCAAATGAATTCCTCA 300
 DB 5332 GCTAGCCAGTTGTCTGCGCTTCTCTGAAAGCGACATTAATCCAAATGAATTCCTCA 5391
 QY 301 GAGCGTGAAGTATGAGGCAACCTCTGTCGCGGCAAGATGAGTGGCGGGAATTAAC 360
 DB 5392 GAGCGTGAAGTATGAGGCAACCTCTGTCGCGGCAAGATGAGTGGCGGGAATTAAC 5451
 QY 361 CGCGTGAAGTATGAGGCAACGATTAATCTGAGCTCTTTCAGCCGCTCCGAGCGGAG 420

DB 5452 CGCGTGAAGTATGAGGCAAAATGATTAATTTTGACTCTTTCGAGCCGCTCCAGCGGAG 5511
 QY 421 GAGGATGACGGGAAATGTCCCTCCCGGAGATCTCTCGGAATTCGAATTTCCCA 480
 DB 5512 GAGGATGACGGGAAATGTCCCTCCCGGAGATCTCTCGGAATTCGAATTTCCCA 5571
 QY 481 CCAGCGATGCCCGCAATGGGACGCGCGGATTAACAACCTCCGCTGTGAGTCTTGAAG 540
 DB 5572 CGAGCGATGCCCGCAATGGGACGCGCGGATTAACAACCTCTCACTTTAAGTCTTGAAG 5631
 QY 541 GCCCGGATGACGCTCCCTCAGTGTATCAATGGGATGCCCACTGCACTTAAGACCCCT 600
 DB 5632 GACCGGATGACGCTCCCTCAGTGTATCAATGGGATGCCCACTGCACTTTAAGTCTTGAAG 5691
 QY 601 CCTATTCACCTTCCACGAGAAAGAGACAGTTGTTCTACAGATTCACCGTGTCTTCT 660
 DB 5692 CCGATTCACCTTCCACGAGAAAGAGACAGTTGTTCTCTGTCAAGAACTCAACGTTCTTCT 5751
 QY 661 GCCCTGAGGAGCTTGGCCACAAAGGCTTTGTTAGTCCGAGACCGTCCGCTCCGACAC 720
 DB 5752 GCTTGGCGGAGCTGCGCACAAAGACCTTGGCAGCTCCGAATGTCGCGCTCCGACAC 5811
 QY 721 GGCACGCAACCGCCCTCTCTGACCAATCTCTCCAGACGAGCGGACAGATCTGACGTT 780
 DB 5812 GGCACGCAACCGCCCTCTCTGACCAAGCTCTCCAGACGAGCGGACAGATCTGACGTT 5871
 QY 781 GAGTGTATTTCTCCATGCCCCCTTTGAGGGGAGCCGGGGACCCCGATCTCAGCGAC 840
 DB 5872 GAGTGTATTTCTCCATGCCCCCTTTGAGGGGAGCCGGGGACCCCGATCTCAGCGAC 5931
 QY 841 GGGTCTGTTCTACCGTGAAGTGAAGAGCGGGGTGAGAGCGTGTGCTGCTGATGTC 900
 DB 5932 GGGTCTGTTCTACCGTGAAGTGAAGAGCGGGGTGAGAGCGTGTGCTGCTGATGTC 5991
 QY 901 TACATGAGACAGCGCTCTGATCAAGCCATGCTGCGGAGAAAGCAAGCTGCCATC 960
 DB 5992 TACATGAGACAGCGCTCTGATCAAGCCATGCTGCGGAGAAAGCAAGCTGCCATC 6051
 QY 961 AACGCTTGAAGCACTTTTGTGCTGCTACCCCAACATGTTTGAAGTTCACATCCGC 1020
 DB 6052 AATGCACTGAGCAACTTTTGTGCTGCTACCCCAACATGTTTGAAGTTCACATCCGC 6111
 QY 1021 AGCGCAAGCCAGCGG 1035
 DB 6112 AGCGCAAGCTGCGG 6126

Search completed: February 19, 2005, 01:26:44
 Job time : 582.412 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2005, 00:21:46 ; Search time 178.108 Seconds
(without alignments)
9508.540 Million cell updates/sec

Title: US-09-664-363-4

Perfect score: 1035
Sequence: 1 ACAGAGTCGATGGGTGCG.....CCGACGCCAGCAGCCG 1035

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
6: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
7: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
8: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
9: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
10: /cgn2_6/prodata/1/ina/5B_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1035	100.0	1035	US-08-191-160-4	Sequence 4, Appl
2	1025.4	99.1	3750	US-08-191-160-22	Sequence 22, Appl
3	885.8	85.6	8642	US-10-029-907-2	Sequence 2, Appl
4	884.6	85.5	7989	US-09-539-601-10	Sequence 10, Appl
5	884.6	85.5	8001	US-09-539-601-7	Sequence 7, Appl
6	884.6	85.5	8637	US-09-539-601-4	Sequence 4, Appl
7	884.6	85.5	8638	US-10-029-907-6	Sequence 6, Appl
8	884.6	85.5	8638	US-10-029-907-24	Sequence 24, Appl
9	884.6	85.5	8639	US-10-029-907-1	Sequence 1, Appl
10	884.6	85.5	8643	US-09-539-601-13	Sequence 13, Appl
11	884.6	85.5	11076	US-09-539-601-1	Sequence 1, Appl
12	883	85.3	8638	US-10-029-907-7	Sequence 7, Appl
13	883	85.3	8638	US-10-029-907-25	Sequence 25, Appl
14	883	85.3	8643	US-10-029-907-4	Sequence 4, Appl
15	883	85.3	8648	US-10-029-907-5	Sequence 5, Appl
16	879.8	85.0	8001	US-09-539-601-16	Sequence 16, Appl
17	879.8	85.0	8001	US-09-539-601-22	Sequence 22, Appl
18	879.8	85.0	11076	US-09-539-601-19	Sequence 19, Appl
19	879.8	85.0	11076	US-09-539-601-25	Sequence 25, Appl
20	878.2	84.9	1176	US-08-537-811-41	Sequence 41, Appl
21	878.2	84.9	9472	US-08-150-204E-96	Sequence 96, Appl
22	876.6	84.7	8001	US-09-539-601-28	Sequence 28, Appl
23	876.6	84.7	11076	US-09-539-601-31	Sequence 31, Appl
24	872.4	84.3	2991	US-08-324-977-49	Sequence 49, Appl
25	872.4	84.3	2991	US-08-384-616-49	Sequence 49, Appl
26	872.4	84.3	2991	US-08-904-686A-49	Sequence 49, Appl
27	872.4	84.3	2991	US-09-315-850-49	Sequence 49, Appl

28	872.4	84.3	7863	1	US-08-324-977-35	Sequence 35, Appl
29	872.4	84.3	7863	2	US-08-384-616-35	Sequence 35, Appl
30	872.4	84.3	7863	3	US-08-904-686A-35	Sequence 35, Appl
31	872.4	84.3	7863	3	US-09-315-850-35	Sequence 35, Appl
32	872.4	84.3	7917	2	US-08-324-977-31	Sequence 31, Appl
33	872.4	84.3	7917	2	US-08-384-616-31	Sequence 31, Appl
34	872.4	84.3	7917	2	US-08-904-686A-31	Sequence 31, Appl
35	872.4	84.3	7917	3	US-09-315-850-31	Sequence 31, Appl
36	872.4	84.3	9030	1	US-08-324-977-13	Sequence 13, Appl
37	872.4	84.3	9030	2	US-08-384-616-13	Sequence 13, Appl
38	872.4	84.3	9030	2	US-08-904-686A-13	Sequence 13, Appl
39	872.4	84.3	9030	3	US-09-315-850-13	Sequence 13, Appl
40	872.4	84.3	9416	1	US-08-324-977-1	Sequence 1, Appl
41	872.4	84.3	9416	2	US-08-384-616-1	Sequence 1, Appl
42	872.4	84.3	9416	2	US-08-904-686A-1	Sequence 1, Appl
43	872.4	84.3	9416	3	US-09-315-850-1	Sequence 1, Appl
44	872.4	84.3	9416	3	US-08-823-895A-27	Sequence 27, Appl
45	862.2	83.3	9595	3	US-09-014-416-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-191-160-4
Sequence 4, Application US/08191160
Patent No. 6210675
GENERAL INFORMATION:
APPLICANT: Highfield, Peter Edmund
APPLICANT: Rodgers, Brian Colin
APPLICANT: Tedder, Richard Seton
APPLICANT: Barbara, John Anthony James
TITLE OF INVENTION: Viral Agent
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESSES:
ADDRESSES: Rothwell, Pigg, Ernst & Kurz
STREET: 1700 K Street
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy diskette, 5.25 inch, 360 Kb storage
COMPUTER: IBM AT compatible
OPERATING SYSTEM: MS-DOS V3.2
SOFTWARE: Wordperfect 5.0 (DOS text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/191,160
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/628,516
FILING DATE: 17 DEC 1990
APPLICATION NUMBER: UK 89 28 562.1
FILING DATE: 18 DEC 1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: UK 90 04 414.0
FILING DATE: 27 FEB 1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: UK 90 04 814.1
FILING DATE: 03 MAR 1990
ATTORNEY/AGENT INFORMATION:
NAME: E. Anthony Figg
REGISTRATION NUMBER: 27,195
REFERENCE/DOCKET NUMBER: 1645-103A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 833-5740
TELEFAX: (202) 833-5744
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1035 base pairs
TYPE: nucleotide with corresponding protein
STRANDEDNESS: single

```

; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to genomic RNA
; ORIGINAL SOURCE:
; ORGANISM: human; serum infectious for PT-NANBH
; IMMEDIATE SOURCE:
; LIBRARY: clone JG3 from cDNA library in lambda gt11
; FEATURE:
; LOCATION: from 1 to 1035 bp portion of the PT-NANBH
; LOCATION: polyprotein
; OTHER INFORMATION: probably encodes viral non-structural
; OTHER INFORMATION: proteins
;
US-08-191-160-4

```

```

Query Match      100.0%; Score 1035; DB 3; Length 1035;
Best Local Similarity 100.0%; Pred. No. 4,3e-276;
Matches 1035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 ACAGAAAGTGAATGGGCTGCGCTGCAACAGGTACGCTCCGCGCTGCAAACTCTCTCAACGG 60
DB 1 ACAGAAAGTGAATGGGCTGCGCTGCAACAGGTACGCTCCGCGCTGCAAACTCTCTCAACGG 60
QY 61 GAGGAGGTGACATTCAGGTCGCGCTCAACCAATACCTGTTGGGTGCGAGCTCCCATGC 120
DB 61 GAGGAGGTGACATTCAGGTCGCGCTCAACCAATACCTGTTGGGTGCGAGCTCCCATGC 120
QY 121 GAGCCCGAACCAGATGATGCTCACTTCCATGCTCAGCCGCTCCCAATCACA 180
DB 121 GAGCCCGAACCAGATGATGCTCACTTCCATGCTCAGCCGCTCCCAATCACA 180
QY 181 GCAGAGACGGCTAAGCGCAGGCTGCGCAAGGGGTCTCCCTCTTGGCCAGCTTTCA 240
DB 181 GCAGAGACGGCTAAGCGCAGGCTGCGCAAGGGGTCTCCCTCTTGGCCAGCTTTCA 240
QY 241 GCTAGCCAGTTGTCGCGCTTCTCGAAGGGACATTAATCCCAATGACTTCCCA 300
DB 241 GCTAGCCAGTTGTCGCGCTTCTCGAAGGGACATTAATCCCAATGACTTCCCA 300
QY 301 GAGCTGACCTCATCGAGCCCACTCTGTGGCGGATGAGATGGGCGGACATTACC 360
DB 301 GAGCTGACCTCATCGAGCCCACTCTGTGGCGGATGAGATGGGCGGACATTACC 360
QY 361 CGGTGAGTCAAGAGCAAGGTATGTAATCTTGGACTCTTTCGACCGCTCCGAGCGAG 420
DB 361 CGGTGAGTCAAGAGCAAGGTATGTAATCTTGGACTCTTTCGACCGCTCCGAGCGAG 420
QY 421 GAGGATGACGGGGAATGTCCTGCGGCGGAGATCTCGGAAATCCAAATTCCTCCA 480
DB 421 GAGGATGACGGGGAATGTCCTGCGGCGGAGATCTCGGAAATCCAAATTCCTCCA 480
QY 481 CCAGCATGCCGATGAGGACGCGCGGATTAACAACCTCCGCTGCTGAGTCTGGAAG 540
DB 481 CCAGCATGCCGATGAGGACGCGCGGATTAACAACCTCCGCTGCTGAGTCTGGAAG 540
QY 541 GCCCGGACCTAGCTCCCTCAAGTGTATGAGTGGTCCCACTGCACTTAAGACCCCT 600
DB 541 GCCCGGACCTAGCTCCCTCAAGTGTATGAGTGGTCCCACTGCACTTAAGACCCCT 600
QY 601 CCTATCACTCTCAAGGAAAGAGAGTGTCTGACGAATCCACCGGTCTTCT 660
DB 601 CCTATCACTCTCAAGGAAAGAGAGTGTCTGACGAATCCACCGGTCTTCT 660
QY 661 GCCCTGGCGAGCTTTCACAAAGCTTTTGTAGTCCGACCGTGGCTGCAACGC 720
DB 661 GCCCTGGCGAGTTCACAAAGCTTTTGTAGTCCGACCGTGGCTGCAACGC 720
QY 721 GGCACGGCAACCGCCCTCTCTGACCAATCTCTCCAGCAGCGGAGAGATCTGAGCTT 780
DB 721 GGCACGGCAACCGCCCTCTCTGACCAATCTCTCCAGCAGCGGAGAGATCTGAGCTT 780
QY 781 GAGTGTATTCCTCAATGCCCCCTTGAAGGGGAGACCGGAGGATCCGATCTCAGCGAC 840
DB 781 GAGTGTATTCCTCAATGCCCCCTTGAAGGGGAGACCGGAGGATCCGATCTCAGCGAC 840

```

```

QY 841 GGGTCTGTCTACCGTAGTAGAGAGCCGGTAGAGACGTGCTGCTGCTCGATGTC 900
DB 841 GGGTCTGTCTACCGTAGTAGAGAGCCGGTAGAGACGTGCTGCTGCTCGATGTC 900
QY 901 TACACATGACAGCGGCTCTGATCAGCCCATGCGCTGGGAGGAAAGCAAGTGGCCATC 960
DB 901 TACACATGACAGCGGCTCTGATCAGCCCATGCGCTGGGAGGAAAGCAAGTGGCCATC 960
QY 961 AACGGTTGAGCAATCTTTGTGCGGTGACCAACAATGCTAGTACCAATCCCGC 1020
DB 961 AACGGTTGAGCAATCTTTGTGCGGTGACCAACAATGCTAGTACCAATCCCGC 1020
QY 1021 AGCGCAAGCCAGCGG 1035
DB 1021 AGCGCAAGCCAGCGG 1035

RESULT 2
US-08-191-160-22
; Sequence 22, Application US/08191160
; Patent No. 6210675
; GENERAL INFORMATION:
; APPLICANT: Highfield, Peter Edmund
; APPLICANT: Rodgers, Brian Colin
; APPLICANT: Tedder, Richard Seton
; APPLICANT: Barbara, John Anthony James
; TITLE OF INVENTION: Viral Agent
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rothwell, Figg, Ernt & Kurz
; STREET: 1700 K Street
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy diskette, 5.25 inch, 360 KB storage
; COMPUTER: IBM AT compatible
; OPERATING SYSTEM: MS-DOS V3.2
; SOFTWARE: Wordperfect 5.0 (DOS text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/191,160
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/628,516
; FILING DATE: 17 DEC 1990
; APPLICATION NUMBER: UK 89 28 562.1
; FILING DATE: 18 DEC 1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 90 04 414.0
; FILING DATE: 27 FEB 1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 90 04 814.1
; FILING DATE: 03 MAR 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: E. Anthony Figg
; REGISTRATION NUMBER: 27,195
; REFERENCE/DOCKET NUMBER: 1645-103A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 833-5740
; TELEFAX: (202) 833-5744
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3750 base pairs
; TYPE: nucleotide with corresponding protein
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to genomic RNA
; ORGANISM: human; serum infectious for PT-NANBH
; IMMEDIATE SOURCE:
; LIBRARY: cDNA clones from 3' end of the genome

```

FEATURE:
LOCATION: from 1 to 3750 bp portion of the PT-NANBH
LOCATION: polypeptide
OTHER INFORMATION: viral non-structural proteins
US-08-191-160-22

Query Match 99.1%; Score 1025.4; DB 3; Length 3750;
Best Local Similarity 99.4%; Pred. No. 3.1e-273;
Matches 1029; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

1 ACAGAGTGGATGGGCTGGCGCTGCAAGTACGCTCCGCGTGCAGAACTCTCTACGG 60
1696 ACAGAGTGGATGGGCTGGCGCTGCAAGTACGCTCCGCGTGCAGAACTCTCTACGG 1755
61 GAGAGGTCACTTCCAGGTGGGCTCAACCAATACCTGGTGGTGCAGCTTCCATGC 120
1756 GAGAGGTCACTTCCAGGTGGGCTCAACCAATACCTGGTGGTGCAGCTTCCATGC 1815
121 GAGCCGGAACGGATATGAGTGTCACTTCCATGCTACCGAACCCCTCCCATCACA 180
1816 GAGCCGGAACGGATATGAGTGTCACTTCCATGCTACCGAACCCCTCCCATCACA 1875
181 GAGAGACGGCTTAAGCGAGGCTGCGCAAGGGGCTCTCCCTCTCTTGGCCAGCTTCA 240
1876 GAGAGACGGCTTAAGCGAGGCTGCGCAAGGGGCTCTCCCTCTCTTGGCCAGCTTCA 1935
241 GCTAGCCAGTGTCTGGCCCTCTCTGAAAGCGACATACCTTCCAAATGACTTCCCA 300
1936 GCTAGCCAGTGTCTGGCCCTCTCTGAAAGCGACATACCTTCCAAATGACTTCCCA 1995
301 GAGCGTGAAGTCAATGAGGCGAACCTCTGTCGGCGAGTGAATGGCGGGGCACTTACC 360
1996 GAGCGTGAAGTCAATGAGGCGAACCTCTGTCGGCGAGTGAATGGCGGGGCACTTACC 2055
361 CGGCTGAGTCAAGAGAAAGAGTATCTCTGACTCTTCCAGCCGCTCCGAGCGAG 420
2056 CGGCTGAGTCAAGAGAAAGAGTATCTCTGACTCTTCCAGCCGCTCCGAGCGAG 2115
421 GAGGATGAGCGGGAAGTGTCCGTCGCGCGAGATCTCGGAAATCCAAAGAAATCCCA 480
2116 GAGGATGAGCGGGAAGTGTCCGTCGCGCGAGATCTCGGAAATCCAAAGAAATCCCA 2175
481 CGAGCATCCCGCATGGGAGCGCCCGGATTAACCTCCGCTGCTGAGTCTGGAG 540
2176 CGAGCATCCCGCATGGGAGCGCCCGGATTAACCTCCGCTGCTGAGTCTGGAG 2235
541 GCGCCGGAAGTACGCTCCCTCAAGTGTATCATGGGTGCCATGCGCACTTAAGACCCCT 600
2236 GCGCCGGAAGTACGCTCCCTCAAGTGTATCATGGGTGCCATGCGCACTTAAGACCCCT 2295
601 CCTATACCACTTCACGAGAAAGAGAGAGTGTCTGACAGAAATCCACCGTCTTCT 660
2296 CCTATACCACTTCACGAGAAAGAGAGAGTGTCTGACAGAAATCCACCGTCTTCT 2355
661 GCGCTGCGGAGTGTCCCAAAAGGCTTTGAGTGTCTGCGGACCGTCCGCTGCACAG 720
2356 GCGCTGCGGAGTGTCCCAAAAGGCTTTGAGTGTCTGCGGACCGTCCGCTGCACAG 2415
721 GCGACGGAACCGCCCTCTCTGACCAATCTCCGACGACGCGGAGAGAGATCTGAGCT 780
2416 GCGACGGAACCGCCCTCTCTGACCAATCTCCGACGACGCGGAGAGAGATCTGAGCT 2475
781 GAGTGTATCTCTCATGCCCCCTTTGAGGGGAGCGCGGGGACCCCGATCTCAGGAC 840
2476 GAGTGTATCTCTCATGCCCCCTTTGAGGGGAGCGCGGGGACCCCGATCTCAGGAC 2535
841 GGGCTTTGGTCAACGAGTGAAGAGGCGGTGAGGAGCGTCTGCTGCTGAGATGTC 900
2536 GGGCTTTGGTCAACGAGTGAAGAGGCGGTGAGGAGCGTCTGCTGCTGAGATGTC 2595
901 TACACATGACAGCGCTGATCAGCCCATGCGCTGCGAGAGAAAGCAAGCTGCCATC 960
2596 TACACATGACAGCGCTGATCAGCCCATGCGCTGCGAGAGAAAGCAAGCTGCCATC 2655

961 AACGGTGAAGAACTTTGGTGGTCAACCAACATGCTTACGCTACCATCCCG 1020
2656 AACGGTGAAGAACTTTGGTGGTCAACCAACATGCTTACGCTACCATCCCG 2715
1021 AGCGCAAGCCAGCGG 1035
2716 AGCGCAAGCCAGCGG 2730

RESULT 3
US-10-029-907-2
Sequence 2, Application US/10029907
Patent No. 6706874

GENERAL INFORMATION:
APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
TITLE OF INVENTION: HEPATITIS C VIRUS
FILE REFERENCE: 13/083
CURRENT APPLICATION NUMBER: US/10/029,907
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 60/257,857
PRIOR FILING DATE: 2000-12-22
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 2

LENGTH: 8642

TYPE: DNA

ORGANISM: HCV

FEATURE:

NAME/KEY: CDS

LOCATION: (1802)... (8407)

NAME/KEY: variation

LOCATION: 6268

OTHER INFORMATION: r = a or g

NAME/KEY: variation

LOCATION: 4446

OTHER INFORMATION: r = a or g

US-10-029-907-2

Query Match 85.6%; Score 885.8; DB 4; Length 8642;
Best Local Similarity 90.9%; Pred. No. 1.6e-234;
Matches 941; Conservative 1; Mismatches .93; Indels 0; Gaps 0;

1 ACAGAGTGGATGGGCTGGCGCTGCAAGTACGCTCCGCGTGCAGAACTCTCTACGG 60
5741 ACAGAGTGGATGGGCTGGCGCTGCAAGTACGCTCCGCGTGCAGAACTCTCTACGG 5800
61 GAGGAGTCACTTCCAGGTCCGCGCTCAACCAATACCTGGTGGTGCAGCTCCCATGC 120
5801 GAGGAGTCACTTCCAGGTCCGCGCTCAACCAATACCTGGTGGTGCAGCTCCCATGC 5860
121 GAGCCGGAACCGGATGATGAGTGTCACTTCCATGCTCAACGACCCCTCCCATCACA 180
5861 GAGCCGGAACCGGATGATGAGTGTCACTTCCATGCTCAACGACCCCTCCCATCACA 5920
181 GCGAGAGCGGCTTAAGCGGAGGCTGCGCAAGGGGCTCTCCCTCTTGGGCACTTCA 240
5921 GCGAGAGCGGCTTAAGCGGAGGCTGCGCAAGGGGCTCTCCCTCTTGGGCACTTCA 5980
241 GCTAGCCAGTGTCTGCGCCCTCTCTGAAAGGAGCAATACCAATGACTTCCCA 300
5981 GCTAGCCAGTGTCTGCGCCCTCTCTGAAAGGAGCAATACCAATGACTTCCCA 6040
301 GAGCGTGAAGTCAATGAGGCGAACCTCTGTCGGCGGATGAGATGGGCGGGGACATTAAC 360
6041 GAGCGTGAAGTCAATGAGGCGAACCTCTGTCGGCGGATGAGATGGGCGGGGACATTAAC 6100
361 CGGCTGAGTCAAGAGAAAGAGTATCTCTGACTCTTCCAGCCGCTCCGAGGAG 420
6101 CGGCTGAGTCAAGAGAAAGAGTATCTCTGACTCTTCCAGCCGCTCCGAGGAG 6160
421 GAGGATGAGCGGGAAGTGTCCGTCGCGCGGAGATCTCGGAAATCCAAAGAAATCCCA 480

```
Db 6161 GAGGATGAGAGGGAAGATATCCGTTCCGGCGGAGATCTCGGAGGTCGACGAAATCCCT 6220
Qy 481 CCAGCGATGCCCCGCAATGGGACAGCCCGGATTACAAACCTCCGCTGTGTGAGTCTGGAA 540
Db 6221 CGAGCGATGCCATATGAGGCGACGCCCGGATTACAAACCTCCGCTGTGTGAGTCTGGAA 6280
Qy 541 GCCCGGACCTAGTCCCTCAGTGTGTACATGGTGTCCCACTGCACTCTTAAGACCCCT 600
Db 6281 GACCGGACCTAGTCCCTCAGTGTGTACATGGTGTCCCACTGCACTCTTAAGACCCCT 6340
Qy 601 CCTATACCACTTCACGAGAAAGAGAGAGATTTTCTGACAGATTCACCGTGTCTTCT 660
Db 6341 CCGATATCACTTCACGAGAAAGAGAGAGATTTTCTGCTGTGTGAGATTCACCGTGTCTTCT 6400
Qy 661 GCCCTGGGAGGAGTTCGCAAAAGGTTTGTGAGTCCGAGACCGTGGCCGCTCCAGAC 720
Db 6401 GCGTTGGCGGAGCTCCGCAAAAGATTCGCGAGCTTCGAAATCGTGGCCGCTCCAGAC 6460
Qy 721 GGCACGCGCAACCGCCCTCTCTGACCAATCTCCGACGAGCGGCGAGCAAGATCTGACGTT 780
Db 6461 GGCACGCGCAACCGCCCTCTCTGACCAAGCCCTCCGAGAGCGGCGAGCGGAGATCCGACGTT 6520
Qy 781 GAGTGTATTTCTTCATGCCCCCTTTGAGGGGAGCCGCGGAGACCCGATCTCAGCGAC 840
Db 6521 GAGTGTATTTCTTCATGCCCCCTTTGAGGGGAGCCGCGGAGATCCGATCTCAGCGAC 6580
Qy 841 GGGTGTGTCTACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
Db 6581 GGGTGTGTCTACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6640
Qy 901 TACACATGACAGCGCTCTGTATCAAGCATGCGCTGCGAGAGAAAGCAAGCTGCCATC 960
Db 6641 TACACATGACAGCGCTCTGTATCAAGCATGCGCTGCGAGAGAAAGCAAGCTGCCATC 6700
Qy 961 AACGCGTTGACCACTTTTGTCTGCTACACACATGCTTACGCTTACCAATCCGC 1020
Db 6701 AATGCACTGAGCAACTTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6760
Qy 1021 AGCGCAAGCGAGCGG 1035
Db 6761 AGCGCAAGCGCTGCGG 6775

RESULT 4
US-09-539-601-10
/ Sequence 10, Application US/09539601C
/ Patent No. 6630343
/ GENERAL INFORMATION:
/ APPLICANT: Barten Schlager, Ralf FW
/ TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
/ FILE REFERENCE: all sequences
/ CURRENT APPLICATION NUMBER: US/09/539,601C
/ CURRENT FILING DATE: 2001-08-30
/ EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
/ EARLIER FILING DATE: 1999-04-03
/ NUMBER OF SEQ ID NOS: 51
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 10
/ LENGTH: 7989
/ TYPE: DNA
/ ORGANISM: Hepatitis C virus
/ FEATURE:
/ NAME/KEY: 5'UTR
/ LOCATION: (1)..(341)
/ OTHER INFORMATION: construct I377/NS3-3'/wt
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (342)..(1181)
/ OTHER INFORMATION: hepatitis C virus core-neomycin phosphotransferase
/ FEATURE:
/ NAME/KEY: fusion protein
/ NAME/KEY: RBS
```

```
/ LOCATION: (1190)..(1800)
/ OTHER INFORMATION: internal ribosome entry site from
/ OTHER INFORMATION: encephalomyocarditis virus
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1801)..(7758)
/ OTHER INFORMATION: hepatitis C virus NS3 - 5B
/ FEATURE:
/ NAME/KEY: 3'UTR
/ LOCATION: (7759)..(7989)
/ PUBLICATION INFORMATION:
/ AUTHORS: Lohmann, Volker
/ AUTHORS: Krner, Frank
/ AUTHORS: Koch, Jan-Oliver
/ AUTHORS: Herian, Ulrike
/ AUTHORS: Theilmann, Lorenz
/ AUTHORS: Barten Schlager, Ralf
/ TITLE: Replication of subgenomic hepatitis c virus RNAs in a
/ TITLE: hepatoma cell line
/ JOURNAL: Science
/ VOLUME: 285
/ PAGES: 110-113
/ DATE: 1999-07-02
/ US-09-539-601-10

Query Match 85.5%; Score 884.6; DB 4; Length 7989;
Best Local Similarity 90.9%; Pred. No. 3.4e-234;
Matches 941; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

Qy 1 ACAGAGTGTATGAGGTGGGCTGACAGTACGCTCCGGCGTGTCAACCTCTCTACCG 60
Db 5092 ACAGAGTGTATGAGGTGGGCTGACAGTACGCTCCAGGTGTCAACCTCTCTACCG 5151
Qy 61 GAGGAGTCAATTCAGAGTGGGCTCAACCAATTCGTTGGTGGTCCAGCTCCATGC 120
Db 5152 GAGGAGTCAATTCAGAGTGGGCTCAATCAATTCGTTGGTGGTCCAGCTCCATGC 5211
Qy 121 GAGCCCGAAGCGATGTAGAGTGTCACTTTCATGTCAACGACCCCTCCACATCACA 180
Db 5212 GAGCCCGAAGCGAGTGTAGAGTGTCACTTTCATGTCAACGACCCCTCCACATCACA 5271
Qy 181 GCAGAGACGCTTAAGCGAGGCTGCGCAGGAGGCTGTCCCTCTTGGCCAGCTTTCA 240
Db 5272 GCAGAGACGCTTAAGCGAGGCTGCGCAGGAGGCTGTCCCTCTTGGCCAGCTTTCA 5331
Qy 241 GCTAGCCAGTGTCTGGGCTTCTCTCGAAGCGACATCAATTAACCAAAATGCTTCCA 300
Db 5332 GCTAGCCAGTGTCTGGGCTTCTCTCGAAGCGACATCAATTAACCAAAATGCTTCCA 5391
Qy 301 GACGCTGACCTATCGAGGCAACCTCTGTGGCGGATGTAGATGGCGGGAATTTACC 360
Db 5392 GACGCTGACCTATCGAGGCAACCTCTGTGGCGGAGATGTAGATGGCGGGAATTTACC 5451
Qy 361 GCGGTGAGTCAAGAGATAGTAATCTGTGACTTTTTCACCCGCTCCGAGCGAG 420
Db 5452 GCGGTGAGTCAAGAGATAGTAATCTGTGACTTTTTCGAGCGCTCCGAGCGAG 5511
Qy 421 GAGATGAGCGGGAAGTGTCCGTCGGCGGAGATCTCGGGAATTCAGAAATTTCCA 480
Db 5512 GAGATGAGCGGGAAGTGTCCGTCGGCGGAGATCTCGGGAATTCAGAAATTTCCA 5571
Qy 481 CCAGCGATGCCCGCATGGGACAGCCCGGATTACAAACCTCCGCTGTGAGTCTGGAAG 540
Db 5572 CCAGCGATGCCCGCATATGGGACAGCCCGGATTACAAACCTCCGCTGTGAGTCTGGAAG 5631
Qy 541 GCCCGGACCTAGTCCCTCAGTGTGTACATGGTGTCCCACTGCACTCTTAAGACCCCT 600
Db 5632 GACCGGACCTAGTCCCTCAGTGTGTACATGGTGTCCCACTGCACTCTTAAGACCCCT 5691
Qy 601 CCTATACCACTTCACGAGAAAGAGAGAGATTTTCTGCTGTGTGAGATTCACCGTGTCTTCT 660
Db 5692 CCGATATCACTTCACGAGAAAGAGAGAGATTTTCTGCTGTGTGAGATTCACCGTGTCTTCT 5751
```


Qy	991	TACACATGAGCAGGGGCTGTATCAAGCATGCGTGGGAGAAAGCAAGCTGCCATTC	960
Dd	6004	TACACATGAGCAGGGCCCTTGATCACGCATGCGTGGGAGAAACAAAGCTGCCATTC	6067
Qy	961	AACGGTTGAGCAACTCTTTGGTGGGTACCACCAACATGCTTACGCTTACCATCCGC	1021
Dd	6064	AATGACATGAGCAACTCTTTGGTGGGTACCACCAACATGCTTACGCTTACCATCTGCG	6122
Qy	1021	AGCGCAAGCCAGCGG	1035
Dd	6124	AGCGCAAGCTGCGG	6138
 RESULT 6 US-09-539-601-4 ; Sequence 4, Application US/09539601C ; Patent No. 6630343 ; GENERAL INFORMATION: ; APPLICANT: Bartenschlager, Ralf FM ; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System ; FILE REFERENCE: all sequences ; CURRENT APPLICATION NUMBER: US/09/539,601C ; EARLIER FILING DATE: 2001-08-30 ; EARLIER APPLICATION NUMBER: 199 15 178-4 GERMANY ; NUMBER OF SEQ ID NOS: 1999-04-03 ; SOFTWARE: Patentin Ver. 2.1 ; SEQ ID NO 4 ; LENGTH: 8637 ; TYPE: DNA ; ORGANISM: Hepatitis C virus ; FEATURE: NAME/KEY: 5'UTR LOCATION: (1)..(341) OTHER INFORMATION: construct I377/NS2-3'/wt FEATURE: NAME/KEY: CDS LOCATION: (342)..(1181) OTHER INFORMATION: HCV core-neomycin phosphotransferase fusion OTHER INFORMATION: protein FEATURE: NAME/KEY: RBS LOCATION: (1190)..(1800) OTHER INFORMATION: internal ribosome entry site from OTHER INFORMATION: encephalomyocarditis virus FEATURE: NAME/KEY: CDS LOCATION: (1801)..(8406) OTHER INFORMATION: hepatitis C virus NS2 - 5B FEATURE: NAME/KEY: 3'UTR LOCATION: (8407)..(8637) PUBLICATION INFORMATION: AUTHORS: Lohmann, Volker AUTHORS: Koch, Jan-Oliver AUTHORS: Herian, Ulrike AUTHORS: Theilmann, Lorenz AUTHORS: Bartenschlager, Ralf TITLE: Replication of subgenomic hepatitis c virus RNAs in a TITLE: hepatoma cell line JOURNAL: Science VOLUME: 285 PAGES: 110-113 DATE: 1999-07-02 US-09-539-601-4			
 Query Match 85.5%; Score 884.6; DB 4; Length 8637; Beet Local Similarity 90.9%; Pred. No. 3.5e-234; Matches 941; Conservative 0; Mismatches 94; Indels 0; Gaps 0			
Qy	1	ACAGAATGATGATGGGTGGCGGTGACACAGATACCTTCGGCGGTGCAAACTCTTCTTAGCG	60

Db	5740	ACAAAGGTGATGGAGGTGCGGTTTCAACAGGTACGCTCAGCGTTCAAACCCCTCTACGG	5799
QY	61	GAGAGGTCACTTTCAGGTGCGGCTCAACCAATACCTGGTGGGTGCACTTCCATGC	120
Db	5800	GAGAGGGTCACTTCTGGTGGGCTCAATCAATACCTGGTGGGTCAAGCTCCATGC	5859
QY	121	GAGCCCGAACCGGATGTATGACAGTGTCTCACTTCATGCTCACCGACCCCTCCCATCA	180
Db	5860	GAGCCCGAACCGGACGTATGACAGTGTCTCACTTCATGCTCACCGACCCCTCCCATACG	5919
QY	181	GCAGAGACGGTAAAGCCGACGGCTGGCCAGGGGGGTCTCCCCCTGCTTGGCCAGCTTTCA	240
Db	5920	GGGAGACGGTAAAGCGGTAGGCTGGCCAGGGGAATCTCCCCCTCTTGGCCAGCTTCA	5979
QY	241	GCTAGCCAGTGTGTGGGCCCTTCTCTCGAAGGCGCATATCACTTACCAGAAATGACTTCCA	300
Db	5980	GCTAGCCAGCTGTGTGGGCCCTTCTCTTGAAGGCATATGACACTACCGCATGACTCCCG	6039
QY	301	GACCTGTACTCATTCAGAGCCCACTCTGTGGCGCATATGATGGCGGGGAATTAACC	360
Db	6040	GACCTGTACTCATTCAGAGCCCACTCTGTGGCGGCAGAGATGGCGGGGAATATCAC	6099
QY	361	CGCGTGGAGTCAGAGAAACAAGTGTATATCTGGACTCTTTCCAGCCGCTCCAGCGGAG	420
Db	6100	CGCGTGGAGTCAGAGAAATATAGGTATGTAATTTTGGACTCTTTTCAGCCGCTCCAGCGGAG	6159
QY	421	GAGATGAGCGGAGAGTGTCCGCTCCGCGGAGATCTGTGGGAATCCAGAAATTTCCA	480
Db	6160	GAGATGAGAGGAGATATCCGTTCCGCGGAGATCTGTGGGAATTTCCCT	6219
QY	481	CCAGCGATGCCGATGGGACAGCCCGGATTAACAACCTCCGCTGCTGAGTCTTGAG	540
Db	6220	CGAGCGATGGCCCATATGGGCAAGCCCGATTAACAACCTCCACTTTATGATCTTGAG	6279
QY	541	GGCCCGGACTACGTCCCTCCAGTGTATCATGGGGCCCACTGGCACTACTTAAGCCCT	600
Db	6280	GACCCGGACTACGTCCCTCCAGTGTATCAAGGATTCATTTGGCCGCTGCCAAAGCCCT	6339
QY	601	CCTATACCACTCCACGAGAGAAAGAGACAGTTGTTCTGCACAGATCCACCCTGCTTCT	660
Db	6340	CCGATACCACTCCACAGGAGAGAGAGCGTGTCTGTGCAGATCTAACGTCCTTCT	6399
QY	661	GCCCTGGCGAGCTTGGCCAAAGGCTTTTGTAAGCTTCGGAACCTGTGGCTGTGACAC	720
Db	6400	GCTTTGGGCGAGCTCGCCACAAAGACCTTCGGCAGCTTCGAAATGTCGGCCGTGACAGC	6459
QY	721	GGCAGCGCAACCGCCCTCGTCAACATCCCTCGACGACGCGCGGAGATCTGAGGTT	780
Db	6460	GGCAGCGCAACCGCTCTCTCTGACACAGCCCTCCGACGACGCGCGGATCTCGAGGTT	6519
QY	781	GAGTGTATTTCTCCATAGCCCCCTTTGAGGGGAGCGGGGGGACCCCGATCTCAGGAC	840
Db	6520	GAGTGTATTTCTCCATAGCCCCCTTTGAGGGGAGCGGGGGATCCCGATCTCAGGAC	6579
QY	841	GGGTCTTGCTTACCGTGAAGTGAAGAGCGCGGTGAGACGTCGTCTGCTGATGTC	900
Db	6580	GGGTCTTGCTTACCGTGAAGTGAAGAGCGGTGAGACGTCGTCTGCTGATGTC	6639
QY	901	TACACATGAGAGGCGCTGTGATACGCACATGCGCTCGAGGAAACGAGTGCATC	960
Db	6640	TACACATGAGAGGCGCTGTGATACGCACATGCGCTCGAGGAAACGAGTGCATC	6699
QY	961	AACCGTTGAGCAACTCTTTGCTCGTGCACCAACAATGTGTACGTATCCATCCGC	1020
Db	6700	AATGCACTGAGCAACTCTTTGCTCGTGCACCAACAATGTGTATCCATCAACTCTGC	6759
QY	1021	AGCGCAAGCCAGCGG 1035	
Db	6760	AGCGCAAGCCAGCGG 6774	

Tue Feb 22 13:23:11 2005

us-09-664-363-4.rn1

Page 7

```
US-10-029-907-6
; Sequence 6, Application US/10029907
; Patent No. 6706874
; GENERAL INFORMATION:
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
; FILE REFERENCE: 13/083
; CURRENT APPLICATION NUMBER: US/10/029,907
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,857
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 8638
; TYPE: DNA
; ORGANISM: HCV
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1802)...(8407)
US-10-029-907-6

Query Match      85.5%; Score 884.6; DB 4; Length 8638;
Best Local Similarity 90.9%; Pred. No. 3.5e-234;
Matches 941; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 1 ACGAAGTGGATGGGTCGCGCTGCACAGTACGCTCGCGCTGCACAACTCTCTACGG 60
DB 5741 ACGAAGTGGATGGGTCGCGCTGCACAGTACGCTCGCGCTGCACAACTCTCTACGG 5800
QY 61 GAGGAGTGCATTCACAGTGGGCTCAACCAATACCTGGTGGTGGAGCTCCCATGC 120
DB 5801 GAGGAGTGCATTCCTGGTGGGCTCAACCAATACCTGGTGGTGGAGCTCCCATGC 5860
QY 121 GAGCCGGAACCGGATGAGAGTGTCTCACTTCCATGCTCACGACCCCTCCCATACA 180
DB 5861 GAGCCGGAACCGGAGTGTGTCTCACTTCCATGCTCACGACCCCTCCCATAGC 5920
QY 181 GCAAGACGGCTAAGGCGAGGCTGGCCAGGGGGTCTCCCTCTCTGGCCAGCTTCA 240
DB 5921 GCGGAGACGGCTAAGGCTGGGCTGGCCAGGGGGTCTCCCTCTCTGGCCAGCTTCA 5980
QY 241 GCTAGCAGTGTCTGGCCCTCTCTGGAAGCGACATTAATCCAAATGACTTCCCA 300
DB 5981 GCTAGCAGTGTCTGGCCCTCTCTGGAAGCGACATTAATCCCTCATGACTCCCG 6040
QY 301 GACGCTGACCTCATCGAGGCAACTCTGTGGCGGATGAGATGGGCGGACATTAC 360
DB 6041 GACGCTGACCTCATCGAGGCAACTCTGTGGCGGAGATGGGCGGGAACATCAC 6100
QY 361 CGCTGAGTCAAGAAACAAGTAGTAATCTGAGTCTTTGACCCGCTCGAGCGGAG 420
DB 6101 CGCTGAGTCAAGAAACAAGTAGTAATTTGAGTCTTTGACCGCTCGCAAGCGAG 6160
QY 421 GAGGATGACGGGGAATGTCCTCCGCGGCGGAGATCTGGGGAATTCAGAAATTTCCA 480
DB 6161 GAGGATGAGGGAATGTCCTCCGCGGCGGAGATCTGGGGAATTCAGAAATTTCCA 6220
QY 481 CCAGCATGCGCGCATGAGGACGCGCGGATTAACAACCTCGCTGAGTCTTGAAG 540
DB 6221 CCAGCATGCGCGCATGAGGACGCGCGGATTAACAACCTCGCTGAGTCTTGAAG 6280
QY 541 GCCCGGACTAGTCCCTCTCAGTGTATCATGGTGCCTGCACTGCACTTAAGACCTT 600
DB 6281 GACCCGCACTAGTCCCTCTCAGTGTATCATGGTGTCTGCACTGCACTTAAGACCTT 6340
QY 601 CCTATACCACTTCAAGGAAAGAGAGACAGTGTCTGACAGATTCACCGGTCTTCT 660
DB 6341 CCGATACCACTTCAAGGAAAGAGAGACAGTGTCTGACAGATTCACCGGTCTTCT 6400
QY 661 GCCCTGCGGAGCTTGCACAAAGCTTTTGTAGTCCGGAACCGTCCGCTGACAGC 720
DB 720
```

```
DB 6401 GCCTTGCGGAGCTCGCCACAAAGACCTTCGAGCTCGGAATCTGCGCCGTCGACAGC 6460
QY 721 GGCACGGCAACCGCCCTCTCTGACCAATCTCTCGAGAGCGCGGACAGATCTGACTT 780
DB 6461 GGCACGGCAACCGCCCTCTCTGACCAATCTCTCGAGAGCGCGGACAGATCTGACTT 6520
QY 781 GAGTCTATTCCTCCATGCGCCCTCTGAGGGGAGCCCGGGAGACCCGATCTCAGCGAC 840
DB 6521 GAGTCTATTCCTCCATGCGCCCTCTGAGGGGAGCCCGGGAGATCCGATCTCAGCGAC 6580
QY 841 GGGCTTGGTCTACCGTGAAGTGAAGAGCGCGGTGAGAGCTGTCTGCTGATGTC 900
DB 6581 GGGCTTGGTCTACCGTGAAGTGAAGAGCGCGGTGAGAGAGCTGTCTGCTGATGTC 6640
QY 901 TACACATGACAGGCGCTGTGATCAGCCATGCGCTGCGGAGGAAGCAAGTCTCCCATC 960
DB 6641 TACACATGACAGGCGCTGTGATCAGCCATGCGCTGCGGAGGAAGCAAGTCTCCCATC 6700
QY 961 AACGCTTGAGCACTTTGTGCTGCTGACCAACATGATGATGATGATGATGATGATG 1020
DB 6701 AATGCACTGAGCACTTTGTGCTGCTGACCAACATGATGATGATGATGATGATG 6760
QY 1021 AGCGAAGCCAGCGG 1035
DB 6761 AGCGAAGCCAGCGG 6775

RESULT 8
US-10-029-907-24
; Sequence 24, Application US/10029907
; Patent No. 6706874
; GENERAL INFORMATION:
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
; FILE REFERENCE: 13/083
; CURRENT APPLICATION NUMBER: US/10/029,907
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,857
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 8638
; TYPE: DNA
; ORGANISM: HCV
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1802)...(8407)
US-10-029-907-24

Query Match      85.5%; Score 884.6; DB 4; Length 8638;
Best Local Similarity 90.9%; Pred. No. 3.5e-234;
Matches 941; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 1 ACGAAGTGGATGGGTCGCGCTGCACAGTACGCTCGCGCTGCACAACTCTCTACGG 60
DB 5741 ACGAAGTGGATGGGTCGCGCTGCACAGTACGCTCGCGCTGCACAACTCTCTACGG 5800
QY 61 GAGGAGTGCATTCACAGTGGGCTCAACCAATACCTGGTGGTGGAGCTCCCATGC 120
DB 5801 GAGGAGTGCATTCCTGGTGGGCTCAACCAATACCTGGTGGTGGAGCTCCCATGC 5860
QY 121 GAGCCGGAACCGGATGAGAGTGTCTCACTTCCATGCTCACGACCCCTCCCATACA 180
DB 5861 GAGCCGGAACCGGAGTGTGTCTCACTTCCATGCTCACGACCCCTCCCATAGC 5920
QY 181 GCAAGACGGCTAAGGCGAGGCTGGCCAGGGGGTCTCCCTCTCTGGCCAGCTTCA 240
DB 5921 GCGGAGACGGCTAAGGCTGGGCTGGCCAGGGGGTCTCCCTCTCTGGCCAGCTTCA 5980
QY 241 GCTAGCAGTGTCTGGCCCTCTCTGGAAGCGACATTAATCCAAATGACTTCCCA 300
DB 300
```

```
Db      5981 GCTAGCCAGCTGTCTGCGCTTCTTGAAGCAATGCACTACCGCTGATGCTCCCG 6040
Qy      301 GACGCTGACCTCATGAGGCAACCTCTGTGCGGCGATGAGTGGCGGGGACATTACC 360
Db      6041 GACGCTGACCTCATGAGGCAACCTCTGTGCGGCGATGAGTGGCGGGGACATTACC 6100
Qy      361 GCGGTGAGTCAAGAAAGGTAGTAACTCTGGACTCTTTGCAACCCGCTCCGAGGAG 420
Db      6101 CCGGTGAGTCAAGAAATAGGTAGTAACTTTGGACTCTTTGAGCCGCTCCAGGCGAG 6160
Qy      421 GAGGATGAGCGGGAAGTGTCCGTCGCGGAGATCTTCGCGAAATCCAAATTTCCA 480
Db      6161 GAGGATGAGCGGGAATATCCGTTCCGCGGAGATCTTCGCGAAATTTCCCT 6220
Qy      481 CAGCGATGCCGATGAGGCAAGCCCGGATTAACAACCTCCGCTGTGAGTCTTGAG 540
Db      6221 CAGCGATGCCATATATGAGGCAAGCCCGGATTAACAACCTCTGTAAGTCTTGAG 6280
Qy      541 GCCCGGAGTACGTCTCCAGTGATACATGAGTGCCCACTGCACTTAAGACCCCT 600
Db      6281 GACCGGAGTACGTCTCCAGTGATACATGAGTGCCCACTGCAAGGCCCT 6340
Qy      601 CCTATACCACTTCAAGGAAAGAGACATGTTCTGACAGAAATCCACCTGTCTTCT 660
Db      6341 CCGATACCACTTCAAGGAAAGAGACATGTTCTGACAGAAATCTACCTGTCTTCT 6400
Qy      661 GCCCTGCGGAGCTTGCACAAAGGCTTTGAGTCTCCGACCGTCCGCGCTGACAG 720
Db      6401 GCTTGTGCGGAGCTGCGCAAAAGACCTTCGCGAGCTTCGAAATCTGCGCGCTGACAG 6460
Qy      721 GGCACGCGCAACGCGCCCTCTGACCAATCTTCGACGACGCGGAGACAGATCTGAC 780
Db      6461 GGCACGCGCAACGCGCTCTCTGACCAAGCCCTCGACGACGCGGAGATCCGACGTT 6520
Qy      781 GAGTGTATCTCTCCATGCCCCCTTGAAGGGGAGACCCGCGATCTCAGCAGC 840
Db      6521 GAGTGTATCTCTCCATGCCCCCTTGAAGGGGAGACCCGCGATCTCAGCAGC 6580
Qy      841 GGGTCTTGGTCTACCGTGAAGGAGCGCGGTGAGAGAGTGTGCTGCTGATGTC 900
Db      6581 GGGTCTTGGTCTACCGTGAAGGAGCGCGGTGAGAGAGTGTGCTGCTGATGTC 6640
Qy      901 TACACATGACAGGCGCTGTGATCAGCGCATGCGCTGCGAGAGAAAGCAAGCTGCCATC 960
Db      6641 TACACATGACAGGCGCTGTGATCAGCGCATGCGCTGCGAGAGAAAGCAAGCTGCCATC 6700
Qy      961 AACGCTTGAAGCACTTTTGTGCTGCTACCAACAATGTCTAGCTACCATCCGC 1020
Db      6701 AATGACTGAGCAACTTTTGTGCTGCTGCTACCAACAATGTCTAGCTACCATCCGC 6760
Qy      1021 AGCGCAAGCGAGCG 1035
Db      6761 AGCGCAAGCGTCCG 6775

RESULT 9
US-10-029-907-1
; Sequence 1, Application US/10029907
; Patent No. 6706874
; GENERAL INFORMATION:
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
; FILE REFERENCE: 13/083
; CURRENT APPLICATION NUMBER: US/10/029,907
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,857
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 8639
; TYPE: DNA
```

```
; ORGANISM: HCV
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1803) ... (8408)
US-10-029-907-1

Query Match      85.5%; Score 884.6; DB 4; Length 8639;
Best Local Similarity 90.9%; Pred. No. 3.5e-234;
Matches 941; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

Qy      1 ACAGAAAGTGAATGGGATGCGGCTGACAGATACGCTCCGGGTGCAAACTCTCTACGG 60
Db      5742 ACAGAAAGTGAATGGGATGCGGCTGACAGATACGCTCCAGGTGCAAACTCTCTACGG 5801
Qy      61 GAGGAGTCAATTCAGAGTGGGCTCAACAATACTGTTGGTGCAGCTCCCATGC 120
Db      5802 GAGGAGTCAATTCAGAGTGGGCTCAATTAATCTGTTGGTGCAGCTCCCATGC 5861
Qy      121 GAGCCGGAACCGAGATGACAGTGTCACTTCATGCTACCGACCCCTCCCAATCA 180
Db      5862 GAGCCGGAACCGAGATGACAGTGTCACTTCATGCTACCGACCCCTCCCAATCA 5921
Qy      181 GCAAGAGCGGCTAAGCGAGGCTGCGCGGGGATCTCCCTCTTGGCCAGCTTTCA 240
Db      5922 GCGGAGAGCGGCTAAGCGAGGCTGCGCGGGGATCTCCCTCTTGGCCAGCTTTCA 5981
Qy      241 GCTAGCCAGTGTCTGCGCTTCTCGAAAGCGACATATACATTAACCAAAATGACTTCCCA 300
Db      5982 GCTAGCCAGTGTCTGCGCTTCTCGAAAGCGACATATACATTAACCAATGACTTCCCA 6041
Qy      301 GACGCTGACCTCATGAGGCGCAACCTCTGTGCGGCAATGAGATGGCGGGGACATTAC 360
Db      6042 GACGCTGACCTCATGAGGCGCAACCTCTGTGCGGCAATGAGATGGCGGGGACATTAC 6101
Qy      361 GCGGTGAGTCAAGAAACAAGTATTAATCTGCACTTTTTCAGACCCGCTCCGAGCGAG 420
Db      6102 GCGGTGAGTCAAGAAACAAGTATTAATCTTTCAGACCCGCTCCGAGCGAG 6161
Qy      421 GAGGATGAGCGGGAAGTGTCCGTCGCGCGGAGATCTCGCGGAAATCCGAAGAAATTTCCCA 480
Db      6162 GAGGATGAGCGGGAAGTATTCGTTTCGCGCGGAGATCTCGCGGAGTCCGAAGAAATTTCCCT 6221
Qy      481 CGAGGATGCCCGCATGAGGCAAGCCCGGATTAACAACCTCTGCTGCTGAGTCTTGAG 540
Db      6222 CGAGGATGCCCATATGAGGCAAGCCCGGATTAACAACCTCTGCTGCTGAGTCTTGAG 6281
Qy      541 GCCCGGAGTACGTCTCTCCAGTGTATATGAGTGTCCCATGCGCATTAATGAAGCCCT 600
Db      6282 GACCGGAGTACGTCTCTCCAGTGTATATGAGTGTCCCATGCGCATTAATGAAGCCCT 6341
Qy      601 CCTATACCACTTCCAGAGAAAGAGACAGTGTCTGACAGAAATCCAGCTGTCTTCT 660
Db      6342 CCGATACCACTTCCAGAGAAAGAGACAGTGTCTGACAGAAATCCAGCTGTCTTCT 6401
Qy      661 GCCCTGCGGAGCTTTCGCAAAAGCTTTTGTGATCTCGAACCGTTCGCTGACAGC 720
Db      6402 GCTTGTGCGGAGCTGCGCAAAAGCTTTCGAGAGCTCCGAATCGTCCGCGTGCAGAGC 6461
Qy      721 GGCAGGCAACCGCCCTCTCTGACCAATCTTCGACGAGCGCGAGAGCATCTGACCTT 780
Db      6462 GGCAGGCAACCGCCCTCTCTGACCAATCTTCGACGAGCGCGAGATCCGACCTT 6521
Qy      781 GAGTGTATCTCTCATGCCCCCTTGAAGGGGAGGCGGGGAGACCCGATCTCAGCAG 840
Db      6522 GAGTGTATCTCTCATGCCCCCTTGAAGGGGAGGCGGGGAGATCCGATCTCAGCAG 6581
Qy      841 GGGTCTTGTCTACCGTGAAGTGAAGAGCGCGTGAAGACGTCTGCTGCTGATGTC 900
Db      6582 GGGTCTTGTCTACCGTGAAGTGAAGAGCGCGTGAAGACGTCTGCTGCTGATGTC 6641
Qy      901 TACACATGACAGGCGCTGTGATCAGCGCATGCGCTGCGAGAGAAAGCAAGTGTGCCATC 960
Db      6642 TACACATGACAGGCGCTGTGATCAGCGCATGCGCTGCGAGAGAAAGCAAGTGTGCCATC 6701
```

Db	5812	GAGGAGGTCAATTCTCTGGTCTGGGCTCAATCAATCACTGGTTGGGTCAAGCTCCATGC	5871
Qy	121	GAGCCCGAACCAGATGTAGCAGTGTCACTTCCATGCTCAACGACCCCTCCCATCA	180
Db	5872	GAGCCCGAACCAGAGTAGCAGTGTCTCACTTCCATGCTCAACGACCCCTCCCATTAACG	5931
Qy	181	GCAAGACGCGTAAGCGCAGGCTGGCCAGGGGGGTCTCCCCCTCTTGGCCAGCTCTTCA	240
Db	5932	GCGGAGACGCGTAAGCGCTAGGCTGGCCAGGGGATCTCCCCCTCTTGGCCAGCTCA	5991
Qy	241	GCTAGCCAGTGTCTGGGCCCTTCTCGAAGGCGCATACATTAACCAAAATGACTTCCCA	300
Db	5992	GCTAGCCAGCTGTCTGGGCCCTTCTTGAAGGCAACATGCACTACCCGTCACTCCCG	6051
Qy	301	GACGCTGACCTCATCGAGGCGCAACCTCTGTGGCGGATGAGATGGCGGGGACATTACC	360
Db	6052	GACGCTGACCTCATCGAGGCGCAACCTCTGTGGCGGAGATGGCGGGGAACATCAACC	6111
Qy	361	CGCGTGAAGTCAGAGAAACAAGTAGTAATCTGTGACTTTCCAGCCGCTCCGAGCGGAG	420
Db	6112	CGCGTGAAGTCAGAGAAATAGTAGTAATTTTGGACTCTTTCGAGCGGCTCCAGCGGAG	6171
Qy	421	GAGGATAGGGGGGAGTCCGCTCCGGCGGAGATCTCGGGGAAATCCAGAAATTTCCA	480
Db	6172	GAGGATAGGGGGGAGTATCCGCTCCGGCGGAGATCTCGGGGAGTCCAGAAATTTCCCT	6231
Qy	481	CCAGCGATGCGCCGATGGGACGCGCCGGAATTACAAACCTCCGCTGAGTCTCTGGAG	540
Db	6232	CGAGCGATGCGCCATATGGGACGCGCCGGAATTACAAACCTCCATCTTAAAGTCTCTGGAG	6291
Qy	541	GCCCCGGACTACGTCCTCCAGATGTATCATGGGTCCCATCTGCCACTACTAAGACCTCT	600
Db	6292	GACCCGGACTACGTCCTCCAGATGTATCAAGGGTGTCAATGGCGCCTGGCAAGGCCCT	6351
Qy	601	CCCTATCACCTCCACGGAAGAAAGGACAGTTGTTGTGACAGAAATCAACGATCTCTCT	660
Db	6352	CCGATACCACTCCACGAGAGAGAGAGAGGATGTCTGTGAGATCTACCGATCTTCT	6411
Qy	661	GCCCTGCGGAGCTTGGCCAAAGGCTTTTGTAGTCCGGAACCGTCGAGCTCGACAGC	720
Db	6412	GCTTGGCGGAGCTCGCCCAAAAGACTTGGGACGCTCCGAATGTGTGGCCGTGACAGC	6471
Qy	721	GGCACGGCAACCGGCCCTCTGACCAATCTCTCGACGAGCGGAGGACGAGATCTGACGTT	780
Db	6472	GGCACGGCAACGGGCTCTCTCTGACCAACCTCTCGACGAGCGGAGATCCGACGTT	6531
Qy	781	GAGTCGTATTTCTTCATATGCCCCCTTGAAGGGGAGAGCCGGGGGACCCCGATTCAGGAC	840
Db	6532	GAGTCGTATTTCTTCATATGCCCCCTTGAAGGGGAGAGCCGGGGGATCCCGATTCAGGAC	6591
Qy	841	GGGCTTGTGTCTACCGTAGTAGAGAGGCGGATGAGACGTGCTGCTGCTGATGTCC	900
Db	6592	GGGCTTGTGTCTACCGTAGTAGAGAGGCTTAGTAGAGACGTGCTGCTGCTGATGTCC	6651
Qy	901	TACACATGACAGGCGCTGTATCAACGCAATGCGCTGGAGAGAAAGCAAGCTGCCATC	960
Db	6652	TACACATGACAGGCGCTGTATCAACGCAATGCGCTGGAGAGAAACCAAGCTGCCATC	6711
Qy	961	AAGCGTTTGAACAATCTTTTGTGCTGCTGACCAACAATGTTTACGCTTACCAATCCCGC	1020
Db	6712	AATGCACTGAGCAACTTTTGTGCTGCTGACCAACAATTTGTCTATGTCTACCAACTTCGC	6771
Qy	1021	AGCGCAAGCGACGG 1035	
Db	6772	AGCGCAAGCGCTGCCG 6786	

```
APPLICANT: Bartschlagel, Ralf FW
FILE REFERENCE: all sequences
CURRENT APPLICATION NUMBER: US/09/539,601C
CURRENT FILING DATE: 2001-08-30
EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
EARLIER FILING DATE: 1999-04-03
NUMBER OF SEQ ID NOS: 51
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 11076
TYPE: DNA
ORGANISM: Hepatitis C virus
FEATURE:
NAME/KEY: 5'UTR
LOCATION: (1)..(341)
OTHER INFORMATION: construct I389/Core-3'/wt
FEATURE:
NAME/KEY: CDS
LOCATION: (342)..(1193)
OTHER INFORMATION: hepatitis C virus core - neomycin
OTHER INFORMATION: phosphotransferase fusion protein
FEATURE:
NAME/KEY: RBS
LOCATION: (1202)..(1812)
OTHER INFORMATION: internal ribosome entry site from
OTHER INFORMATION: encephalomyocarditis virus
FEATURE:
NAME/KEY: CDS
LOCATION: (1813)..(10845)
OTHER INFORMATION: hepatitis C virus polyprotein from core to
OTHER INFORMATION: nonstructural protein NS5B; parental sequence
OTHER INFORMATION: without cell culture-adaptive mutations
FEATURE:
NAME/KEY: 3'UTR
LOCATION: (10846)..(11076)
US-09-539-601-1
```

Query Match 85.5%; Score 884.6; DB 4; Length 11076;
Best Local Similarity 90.9%; Pred. No. 3.8e-234;
Matches 941; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

```
QY 1 ACAGAGTGAATGGGGTGGCGCTGCACAGTACGCTCCGGCGTGCAACCTCTCTACCG 60
DB 8179 ACAGAGTGAATGGGGTGGCGGTGGACAGGTACGCTCCAGCGCAAAACCTCTCTACCG 8238
QY 61 GAGGAGTCACTTCCAGTCCGGCTCAACCAATACCTGGTGGTGGAGCTCCCATC 120
DB 8239 GAGGAGTCACTTCTGGTCCGGCTCAATCAATACCTGGTGGTGGAGCTCCCATC 8298
QY 121 GAGCCCGAAGCGGATGACAGTCACTTCATGCTGACCGAGCCCTCCCAATACCA 180
DB 8299 GAGCCCGAAGCGGATGACAGTCACTTCATGCTGACCGAGCCCTCCCAATACCA 8358
QY 181 GCAGAGACGGCTAAGCGCAGGCTGCGCAGGAGGCTCTCCCGCTCTTGGCCAGCTTTCA 240
DB 8359 GCGGAGACGGCTAAGCGTACGCTGCGCAGGAGATCTCCCGCTCTTGGCCAGCTATCA 8418
QY 241 GCTAGCAGTGTCTGCGCTTCTTGAAGGCGACATTAATCCAAATGACTTTCCA 300
DB 8419 GCTAGCAGTGTCTGCGCTTCTTGAAGGCGACATTAATCCCGTATGACTTCCCG 8478
QY 301 GACGCTGACCTCATCGAGGCGAACCTCTGTGGCGGATGAGATGGGGGGGAGCAATTAC 360
DB 8479 GACGCTGACCTCATCGAGGCGAACCTCTGTGGCGGAGATGGGGGGGAGCAATAC 8538
QY 361 CGCGTGAAGTCAAGAAACAAGTATATCTGACTCTTTTGCACCCGCTCCGAGCGAG 420
DB 8539 CGCGTGAAGTCAAGAAACAAGTATATTTTGGACTCTTTTGCAGCGCTCCAGGCGAG 8598
QY 421 GAGGATGAGCGGGAAGTGTCCGCTCCGCGGAGAGATCTTGGGAAATCCAAATTTCCA 480
DB 8599 GAGGATGAGCGGGAAGTATCCGTTCCGCGGAGATCTTGGGAGTCCAGGAAATTTCCCT 8658
```

```
QY 481 CCAGGATGCCCGATGGGACGCGCCGGATTACAACTCCGCTGTGAGTCTGGAAG 540
DB 8659 CGAGGATGCCCGATATGAGGACGCGCCGGATTACAACTCCAGTGTGAGTCTGGAAG 8718
QY 541 GCCCGGACCTAGTCCCTTCCAGTGTATAGTCCCACTGCGCACTTAAGACCTT 600
DB 8719 GACCCGACCTAGTCCCTTCCAGTGTATAGTCCCACTGCGCGCTTCCAGGCGCTT 8778
QY 601 CCTATACCACTTCCAGGAGAAAGGACATGTTCTTGAAGATTCACCGTCTTCT 660
DB 8779 CGATATACCACTTCCAGGAGAAAGGACATGTTCTTGAAGATTCACCGTCTTCT 8838
QY 661 GCCCTGGGAGACTTCCCAAAAGCTTTGTAGTCTCCGACCGTCCGCGCTGACAGC 720
DB 8839 GCTTGGGAGAGTCCCAAAAGCTTCCGACCGTCCGACCGTCCGCGCTGACAGC 8898
QY 721 GGCACGGCAACCGCCCTCTTGAACCAATCTTCCGACGAGCGGACGAGATCTGACGTT 780
DB 8899 GGCACGGCAACCGCCCTCTTGAACCAATCTTCCGACGAGCGGACGAGATCTGACGTT 8958
QY 781 GAGTGTATTTCTCTCATGCCCCCTTGAAGGAGGACCGGGGACCCCGATCTGACGAC 840
DB 8959 GAGTGTATTTCTCTCATGCCCCCTTGAAGGAGGACCGGGGAGTCCGATCTGACGAC 9018
QY 841 GGGTCTTGTCTACCGTGAAGGAGGCGCGGTGAGAGCTGTCTGTCTGATGTC 900
DB 9019 GGGTCTTGTCTACCGTGAAGGAGGCGCGGTGAGAGCTGTCTGTCTGATGTC 9078
QY 901 TACACATGAGACAGGCGCTGTATCAAGCCATGCTGCGGAGAAAGCAAGCTGCCATC 960
DB 9079 TACACATGAGACAGGCGCTGTATCAAGCCATGCTGCGGAGAAAGCAAGCTGCCATC 9138
QY 961 AACGCTTGAGCAACTCTTTTCTGTGTCACCAACATGCTCTACCAATCCCGC 1020
DB 9139 AATGCACTGAGCAACTCTTTTCTGTGTCACCAACATGCTCTACCAATCCCGC 9198
QY 1021 AGCGCAAGCCAGCGG 1035
DB 9199 AGCGCAAGCCAGCGG 9213
```

RESULT 12
US-10-029-907-7
Sequence 7, Application US/10029907
Patent No. 6706874
GENERAL INFORMATION:
APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
FILE REFERENCE: 13/083
CURRENT APPLICATION NUMBER: US/10/029,907
PRIOR FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 60/257,857
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 8638
TYPE: DNA
ORGANISM: HCV
FEATURE:
NAME/KEY: CDS
LOCATION: (1802)...(8407)
US-10-029-907-7

Query Match 85.3%; Score 883; DB 4; Length 8638;
Best Local Similarity 90.8%; Pred. No. 9.7e-234;
Matches 940; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

```
QY 1 ACAGAGTGAATGGGGTGGCGCTGCACAGTACGCTCCGGCGTGCAAACTCTCTACCG 60
DB 5741 ACAGAGTGAATGGGGTGGCGGTGGACAGGTACGCTCCAGGCTGCAAACTCTCTACCG 5800
```

61 GAGAGAGTCAATTCAGAGTGGGCTCAACCAATACCTGGTGGTGGAGCTCCATGC 120
DB 5801 GAGAGAGTCAATTCAGAGTGGGCTCAACCAATACCTGGTGGTGGAGCTCCATGC 5860
QY 121 GAGCCGGAACCGGATAGAGTGTCACTTCCATGTCTCAACGACCCCTCCCATCACA 180
DB 5861 GAGCCGGAACCGGATAGAGTGTCTCACTTCCATGTCTCAACGACCCCTCCCATCACA 5920
QY 181 GCAGAGACGGCTAAGCGGACGCTGCGCCAGGGGGTCTTCCCTCTCTTGGCAGCTCTTCA 240
DB 5921 GCGGAGACGGCTAAGCGGATAGAGTGTCTGCGCCAGGGGATCTCCCTCTCTTGGCAGCTCTCA 5980
QY 241 GCTAGGCACTGTGTCTGGCCCTTCTCTGGAAGGCAATCACTTCAACCAATGACTTCCCA 300
DB 5981 GCTAGGCACTGTGTCTGGCCCTTCTCTGGAAGGCAATCACTTCAACCAATGACTTCCCA 6040
QY 301 GACGCTGACCTCACTGAGGCGCAACTCTGTGGCGGATGAGATGGGCGGGGACATTAC 360
DB 6041 GACGCTGACCTCACTGAGGCGCAACTCTGTGGCGGATGAGATGGGCGGGGACATTAC 6100
QY 361 CCGGTGAGTCAAGAACAAAGGTATGATCTGTGACTCTTTCGACCCGCTCCGAGCGAG 420
DB 6101 CCGGTGAGTCAAGAACAAAGGTATGATCTGTGACTCTTTCGACCCGCTCCGAGCGAG 6160
QY 421 GAGGATGACGGGAAATGTCTCCCTCCGCGGAGATCTCTGGGAAATTCAGAAATTCCTCA 480
DB 6161 GAGGATGACGGGAAATGTCTCCCTCCGCGGAGATCTCTGGGAAATTCAGAAATTCCTCA 6220
QY 481 CCAGCATGCCCCGATGAGGACGCGCGGATTAACAACCTCCGCTGCTGAGTCTCTGAG 540
DB 6221 CCAGCATGCCCCGATGAGGACGCGCGGATTAACAACCTCCGCTGCTGAGTCTCTGAG 6280
QY 541 GCCCGGACTACGCTCCCTCAAGTGTATGATGGTGGCCCACTGCACTTAAGACCCCT 600
DB 6281 GACCCGGAATACGCTCCCTCAAGTGTATGATGGTGGCCCACTGCACTTAAGACCCCT 6340
QY 601 CCTATACCACTTCAAGGAAAGAGAGACGTTGTCTGACAGATTCACCTGTCTCTT 660
DB 6341 CCGATACCACTTCAAGGAAAGAGAGACGTTGTCTGACAGATTCACCTGTCTCTT 6400
QY 661 GCCCTGGGAGGCTTCCCAAGGCTTTTGTAGTCCGGAACGTCGCGCGCTCGACAG 720
DB 6401 GCCCTGGGAGGCTTCCCAAGGCTTTTGTAGTCCGGAACGTCGCGCGCTCGACAG 6460
QY 721 GGCACGGCAACCGCCCTCTCTGACCAATCTCTCGACGACGCGGAGAGATCTGACGTT 780
DB 6461 GGCACGGCAACCGCCCTCTCTGACCAATCTCTCGACGACGCGGAGAGATCTGACGTT 6520
QY 781 GAGTCTATTTCTTCCATGCCCCCTTGAAGGGGAGCGGGGACCCCGATCTCGACGAC 840
DB 6521 GAGTCTATTTCTTCCATGCCCCCTTGAAGGGGAGCGGGGATCCCGATCTCGACGAC 6580
QY 841 GGGTCTTGGCTTACCGGTGAGTGAAGGAGCGGAGGAGGAGTGTCTGTCTGTCTGATGTC 900
DB 6581 GGGTCTTGGCTTACCGGTGAGTGAAGGAGCGGAGGAGGAGTGTCTGTCTGTCTGATGTC 6640
QY 901 TACACATGACAGGCGCTGTGATCAGCGCAATGCGCTGCGGAGGAAAGCAAGCTGCCATC 960
DB 6641 TACACATGACAGGCGCTGTGATCAGCGCAATGCGCTGCGGAGGAAAGCAAGCTGCCATC 6700
QY 961 AACGCTTGAACAATCTTTTGTCTGCTCACCAACAATGATCTACGATCTCCG 1020
DB 6701 AACGCTTGAACAATCTTTTGTCTGCTCACCAACAATGATCTACGATCTCCG 6760
QY 1021 AGCGCAAGCGAGCGG 1035
DB 6761 AGCGCAAGCGAGCGG 6775

RESULT 13
US-10-029-907-25
; Sequence 25, Application us/10029907

Patent No. 6706874
; GENERAL INFORMATION:
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
; FILE REFERENCE: 13/083
; CURRENT APPLICATION NUMBER: US/10/029,907
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 8638
; TYPE: DNA
; ORGANISM: HCV
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1802)...(8407)
US-10-029-907-25

Query Match 85.3%; Score 883; DB 4; Length 8638;
Best Local Similarity 90.8%; Pred. No. 9,7e-234;
Matches 940; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 1 ACAGAGTGAATGGGATGGGCTGACAGATAGCTCCGCTGCAAACTCTCTTACCG 60
DB 5741 ACAGAGTGAATGGGATGGGCTGACAGATAGCTCCGCTGCAAACTCTCTTACCG 5800
QY 61 GAGAGAGTCAATTCAGAGTGGGCTCAACCAATACCTGGTGGTGGAGCTCCATGC 120
DB 5801 GAGAGAGTCAATTCAGAGTGGGCTCAACCAATACCTGGTGGTGGAGCTCCATGC 5860
QY 121 GAGCCGGAACCGGATAGAGTGTCACTTCCATGTCTCAACGACCCCTCCCATCACA 180
DB 5861 GAGCCGGAACCGGATAGAGTGTCTCACTTCCATGTCTCAACGACCCCTCCCATCACA 5920
QY 181 GCAGAGACGGCTAAGCGGACGCTGCGCCAGGGGGTCTTCCCTCTCTTGGCAGCTCTTCA 240
DB 6161 GAGGATGACGGGAAATGTCTCCCTCCGCGGAGATCTCTGGGAAATTCAGAAATTCCTCA 6220
QY 241 GCTAGGCACTGTGTCTGGCCCTTCTCTGGAAGGCAATCACTTCAACCAATGACTTCCCA 300
DB 5981 GCTAGGCACTGTGTCTGGCCCTTCTCTGGAAGGCAATCACTTCAACCAATGACTTCCCA 6040
QY 301 GACGCTGACCTCACTGAGGCGCAACTCTGTGGCGGATGAGATGGGCGGGGACATTAC 360
DB 6041 GACGCTGACCTCACTGAGGCGCAACTCTGTGGCGGATGAGATGGGCGGGGACATTAC 6100
QY 361 CCGGTGAGTCAAGAACAAAGGTATGATCTGTGACTCTTTCGACCCGCTCCGAGCGAG 420
DB 6221 CCGGTGAGTCAAGAACAAAGGTATGATCTGTGACTCTTTCGACCCGCTCCGAGCGAG 6280
QY 421 GAGGATGACGGGAAATGTCTCCCTCCGCGGAGATCTCTGGGAAATTCAGAAATTCCTCA 480
DB 6161 GAGGATGACGGGAAATGTCTCCCTCCGCGGAGATCTCTGGGAAATTCAGAAATTCCTCA 6220
QY 481 CCAGCATGCCCCGATGAGGACGCGCGGATTAACAACCTCCGCTGCTGAGTCTCTGAG 540
DB 6221 CCAGCATGCCCCGATGAGGACGCGCGGATTAACAACCTCCGCTGCTGAGTCTCTGAG 6280
QY 541 GCCCGGACTACGCTCCCTCAAGTGTATGATGGTGGCCCACTGCACTTAAGACCCCT 600
DB 6281 GACCCGGAATACGCTCCCTCAAGTGTATGATGGTGGCCCACTGCACTTAAGACCCCT 6340
QY 601 CCTATACCACTTCAAGGAAAGAGAGACGTTGTCTGACAGATTCACCTGTCTCTT 660
DB 6341 CCGATACCACTTCAAGGAAAGAGAGACGTTGTCTGACAGATTCACCTGTCTCTT 6400
QY 661 GCCCTGGGAGGCTTCCCAAGGCTTTTGTAGTCCGGAACGTCGCGCGCTCGACAG 720
DB 6401 GCCCTGGGAGGCTTCCCAAGGCTTTTGTAGTCCGGAACGTCGCGCGCTCGACAG 6460

```
QY 721 GGCACGGCAACGCCCTCTCTGACCAATCTCTCGACGACGGGAGACGAGATCTGACGTT 780
DB 6461 GGCACGGCAACGCCCTCTCTCTGACCAATCTCTCGACGACGGGAGATCTGACGTT 6520
QY 781 GAGTCGTATTTCTCCATGCCCCCTTGAAGGGGAGACCCGGGAGATCTCAGCGAC 840
DB 6521 GAGTCGTATTTCTCCATGCCCCCTTGAAGGGGAGACCCGGGAGATCTCAGCGAC 6580
QY 841 GGGTCGTGCTACCGGTGATGAGGAGCGCGGAGAGAGATCTGCTGCTGATGATGTC 900
DB 6581 GGGTCGTGCTACCGGTGATGAGGAGCGCGGAGAGATCTGCTGCTGATGATGTC 6640
QY 901 TACACATGACAGCGCCCTCTGATCAAGCCATGCGCTCGGAGAGAAAGCAAGCTGCCATC 960
DB 6641 TACACATGACAGCGCCCTCTGATCAAGCCATGCGCTCGGAGAGAAAGCAAGCTGCCATC 6700
QY 961 AACGCTTGAGCAACTCTTGTGCTGCTGACACACACATGATGATGATGATGATGATG 1020
DB 6701 AACGCTTGAGCAACTCTTGTGCTGCTGACACACATGATGATGATGATGATGATG 6760
QY 1021 AGCGCAAGCGACGGG 1035
DB 6761 AGCGCAAGCGCTGCGG 6775
```

```
RESULT 14
US-10-029-907-4
; Sequence 4, Application US/10029907
; Patent No. 6706874
; GENERAL INFORMATION:
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
; TITLE OF INVENTION: HEPATITIS C VIRUS
; FILE REFERENCE: 13/083
; CURRENT APPLICATION NUMBER: US/10/029,907
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,857
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 8643
; TYPE: DNA
; ORGANISM: HCV
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1802) ... (8407)
US-10-029-907-4
```

```
Query Match 85.3%; Score 883; DB 4; Length 8643;
Best Local Similarity 90.8%; Pred. No. 9,7e-234;
Matches 940; Conservative 0; Mismatches 95; Indels 0; Gaps 0;
```

```
QY 1 ACAGAAAGTGATGGGGTGGCGGTGCAAGATGAGCTCCGGCGTGCAAACTCTCTTACGG 60
DB 5741 ACAGAAAGTGATGGGGTGGCGGTGCAAGATGAGCTCCGGCGTGCAAACTCTCTTACGG 5800
QY 61 GAGAGAGTCACTTCCAGGTGCGGGCTCAACCAATACCTGGTGGGTCGACGCTCCCATC 120
DB 5801 GAGAGAGTCACTTCCAGGTGCGGGCTCAACCAATACCTGGTGGGTCGACGCTCCCATC 5860
QY 121 GAGCCGGAACCGGATGATGAGAGTCACTTCCATGCTCCGACGACCCCTCCCATCA 180
DB 5861 GAGCCGGAACCGGATGATGAGAGTCACTTCCATGCTCCGACGACCCCTCCCATCA 5920
QY 181 GCAGAGACGGCTAAGCGCAAGCTGCGCAAGGGGCTTCCGCCCTCTTGGCGCAAGCTTCA 240
DB 5921 GCAGAGACGGCTAAGCGCTAAGCGCTGCGCAAGGGGCTTCCGCCCTCTTGGCGCAAGCTTCA 5980
QY 241 GCTAGCCAGTTGCTGCGCTTCTCTGAAAGCGACATCACTTACCCAAAATGACTTCCA 300
DB 5981 GCTAGCCAGTTGCTGCGCTTCTCTGAAAGCGACATCACTTACCCAAAATGACTTCCA 6040
```

```
QY 301 GACGTGACTCTATCGAGGCAACCTCTGTCGGGCAATGAGATGGCGGGACATTAC 360
DB 6041 GACGTGACTCTATCGAGGCAACCTCTGTCGGGCAATGAGATGGCGGGACATTAC 6100
QY 361 GCGGTGAGTCAAGAAACAGATGATCTCTGATCTTTTCCACCCGCTCCGACGGAG 420
DB 6101 GCGGTGAGTCAAGAAACAGATGATCTCTGATCTTTTCCACCCGCTCCGACGGAG 6160
QY 421 GAGATGACGGGAAAGTGTCCGTCGCCGCGAGATCTCTGCGGAATCCAAATAATCCCA 480
DB 6161 GAGATGACGGGAAAGTGTCCGTCGCCGCGAGATCTCTGCGGAATCCAAATAATCCCA 6220
QY 481 CGAGGATGCCCGGATGAGGACGCGCCGATTAACAACCTCCGCTGAGTCTCTGAG 540
DB 6221 CGAGGATGCCCGGATGAGGACGCGCCGATTAACAACCTCCGCTGAGTCTCTGAG 6280
QY 541 GCCCGGACTAGTCTCTCTCCAGTGTATATGGTGTCCCACTGCAACCTTAAGACCCCT 600
DB 6281 GACCGGACTAGTCTCTCTCCAGTGTATATGGTGTCCCACTGCAACCTTAAGACCCCT 6340
QY 601 CCTATACCACTCCACGAGAAAGAGACAGTGTCTGACAGATCCACCGTCTTCT 660
DB 6341 CCTATACCACTCCACGAGAAAGAGACAGTGTCTGACAGATCCACCGTCTTCT 6400
QY 661 GCCCTGGGAGCTTGGCCAAAGGCTTTTGTAGCTCCGACCGCTGCGGACAGC 720
DB 6401 GCCCTGGGAGCTTGGCCAAAGGCTTTTGTAGCTCCGACCGCTGCGGACAGC 6460
QY 721 GGCACGGCAACGCCCTCTCTGACCAATCTCTCCAGACGCGGAGACGATCTGACGTT 780
DB 6461 GGCACGGCAACGCCCTCTCTGACCAATCTCTCCAGACGCGGAGATCTGACGTT 6520
QY 781 GAGTCGTATTTCTCCATGCCCCCTTGAAGGGGAGACCCGGGAGATCTCAGCGAC 840
DB 6521 GAGTCGTATTTCTCCATGCCCCCTTGAAGGGGAGACCCGGGAGATCTCAGCGAC 6580
QY 841 GGGTCGTGCTACCGGTGATGAGGAGCGCGGAGAGATCTGCTGCTGATGATGTC 900
DB 6581 GGGTCGTGCTACCGGTGATGAGGAGCGCGGAGAGATCTGCTGCTGATGATGTC 6640
QY 901 TACACATGACAGCGCCCTCTGATCAAGCCATGCGCTCGGAGAGAAAGCAAGCTGCCATC 960
DB 6641 TACACATGACAGCGCCCTCTGATCAAGCCATGCGCTCGGAGAGAAAGCAAGCTGCCATC 6700
QY 961 AACGCTTGAGCAACTCTTGTGCTGCTGACACACATGATGATGATGATGATGATG 1020
DB 6701 AACGCTTGAGCAACTCTTGTGCTGCTGACACACATGATGATGATGATGATGATG 6760
QY 1021 AGCGCAAGCGACGGG 1035
DB 6761 AGCGCAAGCGCTGCGG 6775
```

```
RESULT 15
US-10-029-907-5
; Sequence 5, Application US/10029907
; Patent No. 6706874
; GENERAL INFORMATION:
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
; TITLE OF INVENTION: HEPATITIS C VIRUS
; FILE REFERENCE: 13/083
; CURRENT APPLICATION NUMBER: US/10/029,907
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,857
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 8648
; TYPE: DNA
; ORGANISM: HCV
; FEATURE:
```

NAME/KEY: CDS
LOCATION: (1802)... (8407)
US-10-029-907-5

Query Match 85.3%; Score 883; DB 4; Length 8648;
Best Local Similarity 90.8%; Pred. No. 9,7e-234;
Matches 940; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

```
QY 1 AAGAGAGTGGAGTGGGCTGGGCTGCAAGGTAGAGCTCCGGCGTGGCAAACTCTCTCCACGG 60
DB 5741 AAGAGAGTGGAGTGGGCTGGGCTGCAAGGTAGAGCTCCGGCGTGGCAAACTCTCTCCACGG 5800
QY 61 GAGAGAGTCAATTCAGAGTGGGCTCAACCAATACCTGGTGGGCTGGAGCTCCATGC 120
DB 5801 GAGAGAGTCAATTCAGAGTGGGCTCAACCAATACCTGGTGGGCTGGAGCTCCATGC 5860
QY 121 GAGCCGGAACCGGATGTAGCAAGTCACTTCCATGCTGACCGACCCCTCCACATCACA 180
DB 5861 GAGCCGGAACCGGATGTAGCAAGTCACTTCCATGCTGACCGACCCCTCCACATCACA 5920
QY 181 GCAAGAGCGGCTAAGCGAGGCTGGCCAGGGGGTCTCCCGCTCTTGGCCAGCTCTCA 240
DB 5921 GCGGAGACCGCTAAGCGAGGCTGGCCAGGGGGTCTCCCGCTCTTGGCCAGCTCTCA 5980
QY 241 GCTAGCCAGTTGCTGGCCCTTCTCGAAAGGCGACATTAATCCAAATGACTTCCCA 300
DB 5981 GCTAGCCAGTTGCTGGCCCTTCTCGAAAGGCGACATTAATCCAAATGACTTCCCA 6040
QY 301 GACGCTGACCTCATCGAGGCCAACCTCTGTGGCGGATGAGATGGGCGGGGACATTAAC 360
DB 6041 GAGGCTGACCTCATCGAGGCCAACCTCTGTGGCGGATGAGATGGGCGGGGACATTAAC 6100
QY 361 CCGGTGAGTCAAGAGAAAGAGTAACTCTGACTCTTTGACCCCGCTCCAGCGGAG 420
DB 6101 CCGGTGAGTCAAGAGAAAGAGTAACTCTGACTCTTTGACCCCGCTCCAGCGGAG 6160
QY 421 GAGGATGAGCGGGAAGTCCGTCGGGGGAGATCTCGGGAATCCAAAGAAATCCCA 480
DB 6161 GAGGATGAGCGGGAAGTCCGTCGGGGGAGATCTCGGGAATCCAAAGAAATCCCA 6220
QY 481 CCAGCGATGCCGATGAGGACCGCCGAGTTACAACTCCGCTGCTGAGTCTGAAAG 540
DB 6221 CCAGCGATGCCGATGAGGACCGCCGAGTTACAACTCCGCTGAGTCTGAAAG 6280
QY 541 GCCCGGAGTACGTCCCTCCAGTGTATACAGGATGCCCACTGCCACCTACTAAGACCCCT 600
DB 6281 GACCCGAGTACGTCCCTCCAGTGTATACAGGATGCCCACTGCCACCCCT 6340
QY 601 CCTATACCACTCCAGGAGAAAGAGACAGTGTCTGACAGATCCACCGGTCTCT 660
DB 6341 CCGATACCACTCCAGGAGAAAGAGACAGTGTCTGACAGATCCACCGGTCTCT 6400
QY 661 GCCCTGGCGAGGCTTCCCAAAAGCTTTTGTAGCTCCGGAACGTCGCGCCGTGACAGC 720
DB 6401 GCTTGGCGGAGGCTTCCCAAAAGCTTCCGGAACGTCGCGCCGTGACAGC 6460
QY 721 GGCACGGCAACCGCCCTCTCTGACCAATCTCCGACGCGCGAGCAGATCTGACGTT 780
DB 6461 GGCACGGCAACCGCCCTCTCTGACCAAGCCTCCGACGCGCGAGATCCGACGTT 6520
QY 781 GAGTCTATTTCTCCATGCCCCCTTGAAGGGGAGACCGGGGATCCCGATCTCAGCGAC 840
DB 6521 GAGTCTATTTCTCCATGCCCCCTTGAAGGGGAGACCGGGGATCCCGATCTCAGCGAC 6580
QY 841 GGGTCTTGGTCTACCGTGAAGGAGGCGCGTGAAGACGTCGTGCTGCTGATGTC 900
DB 6581 GGGTCTTGGTCTACCGTGAAGGAGGCGCTAAGTGAAGACGTCGTGCTGCTGATGTC 6640
QY 901 TACACATGAGCAGGCGCTGTGATCACGCCATGCGCTGCGAGGAGAAAGCAAGCTGCCATC 960
DB 6641 TACACATGAGCAGGCGCTGTGATCACGCCATGCGCTGCGAGGAGAAAGCAAGCTGCCATC 6700
QY 961 AACGCTTGAACACTTTGCTGCGTCAACCAACATGCTTAACGTACCAATCCCGC 1020
```

```
DB 6701 AATGACTAGCAACTTTGCTCCGTACCACTTGTGTATGCTACACATCTCCG 6760
QY 1021 AGCGCAAGCCAGCGG 1035
DB 6761 AGCGCAAGCCAGCGG 6775
```

Search completed: February 19, 2005, 11:51:38
Job time : 181.108 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2005, 00:03:16 ; Search time 3612.18 Seconds
(without alignments)
10906.573 Million cell updates/sec

Title: US-09-664-363-4

Perfect score: 1035
Sequence: 1 ACAGAGTCGATGGGTGGCG.....CCCGCAGCGCAGCCAGCGG 1035

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 14239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	49.2	4.8	925	9	CNS0091P
C 2	45.4	4.4	914	2	BF257968 HVSMEF001
C 3	44.4	4.3	486	4	BG558244 RH122_65
C 4	44.4	4.3	1101	9	CNS0178Y
C 5	44.4	4.3	2100	9	AL108460 Drosophila
C 6	44	4.3	506	5	CL945925 OaIFSB001
C 7	44	4.3	562	7	BP153627
C 8	43.4	4.2	587	6	CN010969
C 9	43	4.2	487	7	CA679773 WIM4.pk00
C 10	43	4.2	540	6	CD727100 4031393 B
C 11	43	4.2	573	6	CD216293 ppp2n.pk0
C 12	43	4.2	591	4	BI391789
C 13	43	4.2	601	4	BM425602 p9f2c.pk0
C 14	43	4.2	615	4	BM425623 p9f2c.pk0
C 15	43	4.2	664	1	AJ450428
C 16	43	4.2	665	1	AJ455557
C 17	43	4.2	667	1	AJ450427
C 18	43	4.2	673	7	CO507218
C 19	43	4.2	674	1	AJ449106
C 20	43	4.2	690	1	AJ445947
C 21	43	4.2	695	4	BM486338
C 22	43	4.2	698	5	BU481178
C 23	43	4.2	704	6	CD763744
C 24	43	4.2	727	6	CD764115

25	43	4.2	737	5	BU281256
26	43	4.2	771	7	CN225007
27	43	4.2	789	7	CN217189
28	43	4.2	830	7	CN232609
29	43	4.2	866	1	AJ447521
30	42.4	4.1	1005	6	CA184993
31	42	4.1	595	1	AI980698
32	41.8	4.0	573	1	AA518979
33	41.8	4.0	584	1	AV922850
34	41.4	4.0	580	1	AV922834
35	41	4.0	673	1	AJ452584
36	41	4.0	628	7	CN140848
37	40.8	3.9	628	9	CL953354
38	40.8	3.9	681	6	CA152295
39	40.8	3.9	828	6	CA141224
40	40.6	3.9	514	7	CN433113
41	40.4	3.9	835	2	BF256551
42	40.2	3.9	420	1	AJ463083
43	40	3.9	509	4	BI996634
44	40	3.9	595	5	BO472474
45	40	3.9	634	5	BO472919

ALIGNMENTS

RESULT 1
CNS0091P/c
LOCUS
DEFINITION
CNS0091P 925 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR19D16 of RPL1-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

ACCESSION
AL053013.1 GI:4934461
VERSION
GSS.
KEYWORDS
Drosophila melanogaster (fruit fly)
SOURCE
Drosophila melanogaster
ORGANISM
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
1 (bases 1 to 925)

AUTHORS
TITLE
JOURNAL

COMMENT

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
Direct Submission
Submitted 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mamoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPL1-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; on bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source
1..925
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone_lib="BACR19D16"
/clone_lib="RPL1-98"
/note="end : TET3"

ORIGIN

Query Match 4.8%; Score 49.2; DB 9; Length 925;
Best Local Similarity 15.6%; Pred. No. 0.082;

Fax: 706 583 0210
Email: mmp@pratt.edu
Seq primer: JEM REV
High quality sequence stop: 345
POLYA=No.

FEATURES

source

Location/Qualifiers
1. 486
/organism="Sorghum propinquum"
/mol_type="rRNA"
/db_xref="taxon:132711"
/note="Organ: Rhizome2 (RHIZ2)"
/note="Organ: Rhizomes; Vector: pBluescript II from Lambda
Zap II; Site 1: XhoI; Site 2: EcoRI; The library was made
from poly-A RNA in the cloning vector lambda Zap II.
Clones to be sequenced were prepared by mass excision."

ORIGIN

Query Match 4.3%; Score 44.4; DB 4; Length 486;
Best Local Similarity 54.2%; Pred. No. 1.4;
Matches 90; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
QY 720 CCGACGCGCAACCGCCCTCTTCAATCTTCCGACGACGCGAGAGATCTGACT 779
DB 22 CCGTACCGCCACCGCATGCGCCGACCATCCGCGCGCGCCGCTCAACCGAGTGA 81
QY 780 TGAAGTGTATTCCTCATGCGCCCTTGAAGGCGGAGCGCGGAGCCCGATCTCAGCA 839
DB 82 CCGGCGCGCTGCGCTCTTCCGACCGCTGCGCGAGAGCGCGCGCGCGCATGGA 141
QY 840 CGGCTCTTGTCTTACCGCTGAGAGAGCGCGCTGAGACGTCTG 885
DB 142 GGAAGCGCGCGCGCGCGCGAGAGGTGCGCGGTGCGCGCGGAC 187

RESULT 4
LOCUS CNS0175Y 1101 bp DNA linear GSS 26-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence Sp6 end of BAC
BACN37L08 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

ACCESSION AL108460.1 GI:5628764
VERSION AL108460
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster

REFERENCE Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)

REFERENCE Direct Submision
TITLE Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
JOURNAL Web : www.genoscope.cns.fr

COMMENT

Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billand at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
Project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.

FEATURES

source

Location/Qualifiers
1. 1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone_lib="BACN37L08"
/clone_1ib="DrosBAC"
/plasmid="pBelobAC11"
/note="end : Sp6"

ORIGIN

Query Match 4.3%; Score 44.4; DB 9; Length 1101;
Best Local Similarity 14.8%; Pred. No. 1.5;
Matches 61; Conservative 176; Mismatches 175; Indels 0; Gaps 0;
QY 61 GAGAGTCACATTCAGGTCGGCTCAACCAATACCTGTTGGCGAGCTCCATGC 120
DB 675 GKKAKGVGRGIVCCAGGAGACACMAADCGCCACMAACSSSSSSSCASTSSAS 734
QY 121 GAGCCCGAACCGGATTTACAGTGTCTCATCTTCATCTCAACGCCCTCCCATCACA 180
DB 735 RGGVSSCACSSBSGGAACAGAGGSGGGGCGCCASSGVCAACASSCSSASMC 794
QY 181 GCAGAGACCGCTTAAGGCGAGGCTGCGCCAGGGGTCTCCCTCTCTTGGCCAGCTTCA 240
DB 795 GSSGSSCCASCGSCCGTGVSSCGAVSASASVMSKVASAVACSAVASGMSAGVSSGR 854
QY 241 GCTAGCAGTGTCTGCGCCCTTCTCGAAGCGGACATCATTAACCAATGACTTCCA 300
DB 855 SSVASSSVASAAVSSSSSSSSSSSVVAAVASSSSSSASMAVAAVAVSVSVAS 914
QY 301 GACGTCAGCTATCGAGCCCAACCTCTGTGCGCGCATGAGATGGCGGACATTAC 360
DB 915 VVSSSSCSSSSASVAVASVAAASVSSSSSSSVTSASVAVSVAVMSAVVSS 974
QY 361 CGGCTGAGTCAGAGAACAGGTAATCTCTGACTCTTTCGACCGCTCGAGCGGAG 420
DB 975 ASSSSSVSVSVVAVASAAAAAASSSSSASAVAVASSSSSSSSSSSAS 1034
QY 421 GAGATGAGCGGGAAGTGTCTCCGCGGAGATCTTCGGAATCCAGA 472
DB 1035 SVSSVSSSCSVSVSVSSSVAVASASVSVAAVAAVAAAAASA 1086

RESULT 5
LOCUS CL945925/c 2100 bp DNA linear GSS 21-SEP-2004
DEFINITION OIRFSB001364 Oryza sativa Express Library Oryza sativa (indica
cultivar-group) genomic, genomic survey sequence.

ACCESSION CL945925.1 GI:52357934
VERSION CL945925
KEYWORDS GSS.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 2100)

REFERENCE Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
Wong, G. K. S., Deng, X. W. and Wang, J.
An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
Unpublished (2004)

COMMENT

Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-8048676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.

FEATURES

source

Location/Qualifiers
1. 2100
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Express Library"
/note="Oryza sativa exon trapped genomic sequences"

ORIGIN

Query Match 4.3%; Score 44.4; DB 9; Length 2100;
Best Local Similarity 59.5%; Pred. No. 1.7;

Matches	75;	Conservative	0;	Mismatches	51;	Indels	0;	Gaps	0;
Qy	639	GACGAGATCCACCGTGTCTTCTGCGCCCTTGCGGAGACTTCCCAAGAGCTTTTGTAAGTC	698						
Db	174	GACGGGAAACCAACCGCCGCTCGGCGCCCTTGCGGAGACTTCCCAAGAGCTTTTGTAAGTC	115						
Qy	699	CGGACCGCGGAGCGGTGCGAGCGGAGCGGCAACCGCCCTCTGACCAATCTCCGACGA	758						
Db	114	CCTCGGAGATGCGGTGCGGCGGCGCCAGCGGAGATCCCGCGCGCCCGCCACCACTCGGCTG	55						
Qy	759	CGGCGG 764							
Db	54	CGGCGG 49							
RESULT 6									
BP153627/c									
LOCUS									
DEFINITION									
ACCESSION									
VERSION									
KEYWORDS									
SOURCE									
ORGANISM									
REFERENCE									
AUTHORS									
TITLE									
JOURNAL									
COMMENT									
FEATURES									
SOURCE									
ORIGIN									
Query Match	4.3%;	Score 44;	DB 5;	Length 506;					
Best Local Similarity	48.1%;	Pred. No. 1.8;							
Matches 125;	Conservative	0;	Mismatches 135;	Indels 0;	Gaps 0;				
539	AGGCCCCGAGCTAGCTCCCTCAGGATGATACAGGAGTGGCCACTGCACTTACTAAGACC	598							
Db	389	AGCGCTGGCTGGCTCCCTCAGGAGCTCTTGACACTTACTGCTATGATTAACCG	330						
Qy	599	CTCCTATACCACTTCACGAGAAAGAGACAGTGTCTTGACAGAAATCCACCGTGTCTT	658						
Db	329	TGCTGCTCGGCGCGAACCGGCTCCGGGAGACGAGGCCAACGCCCGCGCTTTTCCCG	270						
Qy	659	CTGCCCTTGCGGAGCTTCCCAAGAGGCTTTTGTAAGTCTCGGACCGTGGCGGTGACA	718						

Db	QY	Db	QY	Db	QY	RESULT 7
269	CAGCGAGGCGCCGAGCTGCTCCGGAAGCAGCAGGACCGAGCCGCGGAGCCAGCGCGCCT	210				
719	GCGGACGCGCACCCGCTCTTGACCAATCTTCGACGACGCGGAGCAGATCTGAC	778				
209	CGGGTCGCGACGAGGAGGCGCTGACAGGCGCCCAAGCGGCGGCGACCGGGTCCGCG	150				
779	TTGAGTCGTATTCCTCCAG	798				
149	GGGGCGCGAGGCGCCACG	130				
LOCUS	CN010969/c					
DEFINITION	CN010969	652 bp	mrna	linear	EST 29-MAR-2004	
ACCESSION	WHE3878_D02_H04ZS					
VERSION	library Triticum aestivum cDNA clone WHE3878_D02_H04					
KEYWORDS	sequence.					
ORGANISM	CN010969					
SOURCE	CN010969.1	GI:45801001				
REFERENCE	EST.					
AUTHORS	Triticum aestivum (bread wheat)					
	Triticum aestivum					
	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;					
	Poideae; Triticaceae; Triticum.					
	1 (bases 1 to 652)					
	Anderson,O.D., Chao,S., Han,P.S., Heinen,S., Hsia,C.C., Kang,Y.,					
	Kruger,W.M., Lazo,G.R., Miller,S., Muehlbauer,G.J., Miller,R.,					
	Pritsch,C., Rausch,C.J., Seaton,C.L., Tong,J.C., Vance,C. and					
	Wilson,C.F.					
TITLE	The structure and function of the expressed portion of the wheat					
JOURNAL	genomes - Fusarium graminearum infected spike cDNA library					
COMMENT	Unpublished (2001)					
	Contact: Olin Anderson					
	US Department of Agriculture, Agriculture Research Service, Pacific					
	West Area, Western Regional Research Center					
	800 Buchanan Street, Albany, CA 94710, USA					
	Tel: 5105959773					
	Fax: 510595818					
	Email: candersn@pw.usda.gov					
	Sequences have been trimmed to remove vector sequence and low					
	quality sequence with phred score less than 20. No effort was taken					
	to identify ESTs of fungal origin from this library, thus this EST					
	could be of wheat or fungal origin.					
	Seq primer: SK primer.					
FEATURES	Location/Qualifiers					
SOURCE	1..652					
	/organism="Triticum aestivum"					
	/molecule="mrna"					
	/culivar="Sumai3"					
	/db_xref="taxon:4565"					
	/clone="WHE3878_D02_H04"					
	/tissue_type="Spike"					
	/dev_stage="Adult plant"					
	/lab_host="E. coli SOLR"					
	/clone_lib="Wheat Fusarium graminearum infected spike cDNA					
	library"					
	/notes="Vector: Lambda Uni-ZAP XR, excised phagemid					
	plasmid SK, Site_1: EcoRI, Site_2: XhoI; Plants were					
	grown in the greenhouse. Spikes were sprayed at anthesis					
	with Fusarium graminearum. Total RNA, and poly(A) RNA were					
	prepared and pooled from infected spike at 0, 6, 12, 24,					
	36 and 48 hours after inoculation, a cDNA library was					
	made, and the cDNA clones were in vivo excised to give					
	phagescript phagemids in G. Muehlbauer lab at the					
	University of Minnesota (Kruger, W.M., Muehlbauer, G.J.,					
	Pritsch, C., Vance, C.). The cDNA library should contain					
	genes of both wheat and fungal pathogen origin. Plasmid					
	DNA preparations and DNA sequencing were performed in the					
	OD Anderson lab (all other authors)."					

Query Match 4.3%; Score 44; DB 7; Length 652;
 Best Local Similarity 48.8%; Pred. No. 1.8; Indels 0; Gaps 0;
 Matches 119; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

QY 662 CCGTGGCGAGCTTGGCCACAAAGGCTTTTGGTAGCTCCGACCGGTGGCCGTGACAGCG 721
 DB 270 CCGTGGCGAGCGAGCGCCGAGTACATCCGCGCGCGGTAGCTCCGCGCTTTTCCG 211

QY 722 GCGACGCAACCGCCCTCTCTGACCAATCTCTCCAGCAGCGCGGAGCAGATCTGACGTTG 781
 DB 210 CCGTGGCGACCGCCACCGCCGACCGTCTCTCCGCGCGGTAGCGCGCGCGCG 151

QY 782 AGTCGATTCCTTCATGCCCGCTTGGAGGGAGGCGCGGAGCCCGATCTGACGAGC 841
 DB 150 TAACTTCACCGCTTCCCGACCGTGGCCAGGGTAGCCGACCGCGCGCTGCTCCA 91

QY 842 GGTCTTGTCTACCGTGAAGTAGAGAGCGGTGAGAGACGTCTGTGCTGCTGATGCT 901
 DB 90 CCGCTTCCCGACCGTGGCGAGGGTAGCGCGCACCGCGCGCATAGCTCTCCGCGCTTCC 31

QY 902 ACAC 905
 DB 30 CCAC 27

RESULT 8
 CA679773 587 bp mRNA linear EST 24-NOV-2002
 LOCUS wlm4.pk0024.c7 wlm4 Triticum aestivum cDNA clone wlm4.pk0024.c7 5'
 DEFINITION end, mRNA sequence.
 ACCESSION CA679773
 VERSION CA679773.1 GI:25264502
 KEYWORDS EST.
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Triticum.
 REFERENCE 1 (bases 1 to 587)
 Tingey, S.V., Powell, W., Wolters, P., Dolan, M., Hainey, C., Yuan, Z.,
 Miao, G., Caraher, N. and Hanafey, M.K.
 DuPont Wheat cDNA Sequence
 Unpublished (2002)
 CONTACT Scott V. Tingey
 CROP Genetics
 E. I. DuPont de Nemours and Company
 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
 Tel: 302-631-2602
 Fax: 302-631-2607
 Email: Scott.V.Tingey@USA.dupont.com
 Seq primer: M13.
 Location/Qualifiers
 1..587
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /cultivar="Stephens"
 /db_xref="taxon:4565"
 /clone="wlm4.pk0024.c7"
 /issue_type="leaf"
 /clone_lib="wlm4"
 /note="Vector: pBluescript SK+; Site_1: EcoRI; Site_2:
 XhoI; Wheat (Triticum aestivum L.) seedlings 4 hr after
 inoculation w/E. graminis"

ORIGIN
 Query Match 4.2%; Score 43.4; DB 6; Length 587;
 Best Local Similarity 53.3%; Pred. No. 2.6; Indels 0; Gaps 0;
 Matches 89; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 445 CCGGCGAAGTCTTGGCGAAATCAAGAAATTCACACGAGATGCCGAGTGGGACGC 504
 DB 217 CCGCGCGCGCTCTCCGAAACCTTAATCTCGCTCCGCGCGCGTCTCTCCGCG 276

QY 505 CCGATTACCAACCCCTCCGCTGTGAGTCTGTGAAGCCCGGACTACGCTCCGAGT 564
 DB 277 CCGGCGTCCGTATGCTCCGCTCTTGTATTCGGGAAAGCGCTCTTGTCCGCGCAAC 336

QY 565 GTACATGAGGTGCCACTGCGACCTACTAAGACCCCTCTAATACAC 611
 DB 337 TGGCAAGACCAACGCTCTTAATTAAAGCGCAATACATCAAC 383

RESULT 9
 CV039737 487 bp mRNA linear EST 01-OCT-2004
 LOCUS 4137143 BARC 3GAL chicken mixed tissue Gallus gallus cDNA clone
 DEFINITION 3GAL.25G10 5', mRNA sequence.
 ACCESSION CV039737
 VERSION CV039737.1 GI:53559176
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 REFERENCE 1 (bases 1 to 487)
 Evock-Clover, C.M., Ashwell, C.M., McMurtry, J.P., Lillehoj, H.S.,
 Matukumalli, L.K. and Van Tassel, C.P.
 Characterization of expressed sequence tags generated from multiple
 chicken tissues
 Unpublished (2004)
 CONTACT Christina M. Clover
 Growth Biology Laboratory
 Animal and Natural Resources Institute
 Bldg. 200 Rm215 BARC-East, Beltsville, MD 20705, USA
 Tel: 3015048224
 Fax: 3015048623
 Email: chris@anri.barc.usda.gov
 Single pass sequencing. Bases called and trimmed with phred
 0.000925 using options -trim alt '- -trim fastavector identified by
 cross_match using options -mismatch 12 -mnscore 12
 Plate: 25 row: G column: 10
 Seq primer: CCCAGTCACGACGTGTAAACG
 High quality sequence stop: 487.
 Location/Qualifiers
 1..487
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="Leghorn and broiler"
 /db_xref="taxon:9031"
 /clone="3GAL.25G10"
 /lab_host="DH5alpha"
 /clone_lib="BARC_3GAL chicken mixed tissue"
 /note="Vector: pBluescript, SK+, Stratagene; Site_1: NotI;
 Site_2: EcoRI; Normalized library of pooled RNA isolated
 from whole brain, ultimobranchial gland, parathyroid
 gland, cecal tonsil and primordial germ cells Multiple"

ORIGIN
 Query Match 4.2%; Score 43; DB 7; Length 487;
 Best Local Similarity 55.8%; Pred. No. 3.3; Indels 0; Gaps 0;
 Matches 82; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 318 GGGCAACCTCTGTGGCGGATGAGATGGCGGGGACATTAACCGGTGAGTCAAGAA 377
 DB 82 GGGGACCGCTCGACCGGAGGTGAGGCGCGCGCTTAACCGGAGCGCTAGGC 141

QY 378 CAAGTAGTATCTTGACTCTTTCGACCGCTCCGACGAGAGAGATGAGCGGAGT 437
 DB 142 AAGAGTGGATCCCTTAACCGCGCATCAAGCTGTGTAAGAGAGAGCTGACCGGCGCA 201

QY 438 GTCCGTCGCGGCGAGATCTTGGGAA 464
 DB 202 GGAGCGCTTCGCGACCGCTCTGACAGAA 228

JOURNAL
COMMENT

Library USDA/IFAFS Animal Genome Project
Unpublished (2001)
Contact: Larry A. Cogburn
University of Delaware
Townsend Hall, Newark, DE 19717, USA
Tel: 302-831-1335
Fax: 302-831-2822
Email: cogburn@udel.edu, www.chickest.udel.edu.

FEATURES

source

1. .591
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Commercial broiler chicken"
/db_xref="taxon:9031"
/clone="pgfp1.pk005.i7"
/sex="Male and Female"
/tissue_type="Pituitary Gland/Hypothalamus/Pineal Gland"
/dev_stage="Embryonic (dl2,dl4,dl9); post-hatch
(w3,w5,w7,w9)"
/lab_host="E. coli EMDH10B"
/clone_lib="Normalized Chicken
Pituitary/Hypothalamus/Pineal Library"
/note="Vector: PCWSPORT6; Library made from equivalent
pools of total RNA isolated from each tissue at different
ages. Single pass sequencing from 5'-end"

ORIGIN

Query Match

Best Local Similarity 55.8%; Pred. No. 3.3;
Matches 82; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 318 GGGCAACCTCTCTGCGCGCATGATGGCGGAGCATTTACCCGCTGAGTCAGAGAA 377
DB 146 GGGGACCCCGCTGACCGGAGGTGAGGCGCGGACCGGACCGGAGCGCTGAGGC 205
QY 378 CAAGTAGTATCTCTGACTCTTTGACCCGCTCCGAGGAGGAGATGAGCGGGAAT 437
DB 206 AGAGTGGCATCTCTTAACCGCGCATCCAGCTGTGATGAGAGAGAGCTGAGCGGCCCA 265
QY 438 GTCCGTCCCGCGGAGATCTCTGCGAA 464
DB 266 GGAGCGCTCTGCGACCGCTCTGAGAA 292

RESULT 13

LOCUS

DEFINITION

BM425602 601 bp mRNA linear EST 30-JAN-2002
pgf2c.pk001.e8 Primary Chicken Abdominal Fat Library (pgf2c) Gallus
gallus CDNA clone pgf2c.pk001.e8 5' similar to p1r167849|167849
tropomyosin - rat gbl|AA42263.1| (U24776) tropomyosin [Rattus
norvegicus], mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

BM425602
EST.
Gallus gallus (chicken)
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 601)
Cogburn, L.A., Morgan, R. and Burnside, J.
ESTs from Primary Chicken Fat CDNA Library-USDA/IFAFS Animal Genome
Project
Unpublished (2002)
Contact: Larry A. Cogburn
University of Delaware
Townsend Hall, Newark, DE 19717, USA
Tel: 302-831-1335
Fax: 302-831-2822
Email: cogburn@udel.edu, www.chickest.udel.edu.

FEATURES

source

1. .601
/organism="Gallus gallus"

/mol_type="mRNA"
/strain="Commercial broiler, Ottawa Research Centre,
leghorn"
/db_xref="taxon:9031"
/clone="pgf2c.pk001.e8"
/sex="Male and Female"
/tissue_type="Abdominal Fat"
/dev_stage="Embryonic (dl8,dl9); post-hatch
(dl,w3,w5,w7,w9,w16,lyr)"
/lab_host="E. coli EMDH10B"
/clone_lib="Primary Chicken Abdominal Fat Library (pgf2c)"
/note="Vector: PCWSPORT6; Library made from equivalent
pools of total RNA isolated from each developmental age
(across strains); Single pass sequencing from 5'-end"

ORIGIN

Query Match

Best Local Similarity 4.2%; Score 43; DB 4; Length 601;
Matches 82; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 318 GGGCAACCTCTCTGCGCGCATGATGGCGGAGCATTTACCCGCTGAGTCAGAGAA 377
DB 50 GGGGACCCCGCTGACCGGAGGTGAGGCGCGGACCGGACCGGAGCGCTGAGGC 109
QY 378 CAAGTAGTATCTCTGACTCTTTGACCCGCTCCGAGGAGGAGATGAGCGGGAAT 437
DB 110 AGAGTGGCATCTCTTAACCGCGCATCCAGCTGTGATGAGAGAGAGCTGAGCGGCCCA 169
QY 438 GTCCGTCCCGCGGAGATCTCTGCGAA 464
DB 170 GGAGCGCTCTGCGACCGCTCTGAGAA 196

RESULT 14

LOCUS

DEFINITION

BM425623 615 bp mRNA linear EST 30-JAN-2002
pgf2c.pk001.f7 Primary Chicken Abdominal Fat Library (pgf2c) Gallus
gallus CDNA clone pgf2c.pk001.f7 5' similar to p1r167849|167849
tropomyosin - rat gbl|AA42263.1| (U24776) tropomyosin [Rattus
norvegicus], mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

BM425623
EST.
Gallus gallus (chicken)
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 615)
Cogburn, L.A., Morgan, R. and Burnside, J.
ESTs from Primary Chicken Fat CDNA Library-USDA/IFAFS Animal Genome
Project
Unpublished (2002)
Contact: Larry A. Cogburn
University of Delaware
Townsend Hall, Newark, DE 19717, USA
Tel: 302-831-1335
Fax: 302-831-2822
Email: cogburn@udel.edu, www.chickest.udel.edu.

FEATURES

source

1. .615
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Commercial broiler, Ottawa Research Centre,
leghorn"
/db_xref="taxon:9031"
/clone="pgf2c.pk001.f7"
/sex="Male and Female"
/tissue_type="Abdominal Fat"
/dev_stage="Embryonic (dl8,dl9); post-hatch
(dl,w3,w5,w7,w9,w16,lyr)"
/lab_host="E. coli EMDH10B"
/clone_lib="Primary Chicken Abdominal Fat Library (pgf2c)"



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 144962

TO: Bao-Qun Li
Location: rem/3d24/3c18
Art Unit: 1648
Tuesday, February 22, 2005

Case Serial Number: 09/664363

From: Edward Hart
Location: Biotech-Chem Library
REM-1A55
Phone: 571-272-2512

edward.hart@uspto.gov

Search Notes

Examiner Li,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

This Page Blank (uspto)

STIC-Biotech/ChemLib

144962

mg

From: Li, Bao-Qun
Sent: Friday, February 11, 2005 8:08 AM
To: STIC-Biotech/ChemLib

PLEASE DO THE SEQUENCE HOMOLOGY AND INTERFERENCE OF SEQ ID NO; 3 IN APPLICATION SN.
09,664,363. THANKS.

Bao Qun Li M.D
TC 1600
Art Unit 1648
Tel. 517-272-0904
REM, 3C18
Rm. 3D24

RECEIVED
FEB 11 2005
STIC

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2- _____
Date Searcher Picked up: 2/16/05
Date Completed: 2/16/05
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search /

NA Sequence: # _____
AA Sequence: # _____
Structure: # _____
Bibliographic: _____
Litigation: _____
Patent Family: _____
Other: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: PSA
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)



STIC SEARCH RESULTS

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact:*

Mary Hale, Information Branch Supervisor
Remsen Bldg. 01 D86
571-272-2507

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC-Biotech-Chem Library Remsen Bldg.



This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2005, 21:00:40 ; Search time 7668.71 Seconds
(without alignments)
10899.589 Million cell updates/sec

Title: US-09-664-363-3

Perfect score: 1770
Sequence: 1 CAAATGACTTCCCAAGCC.....ACTACTCACCCTGACCG 1770

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ha:*
2: gb_hng:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sbs:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	1770	100.0	1770	6 A28131	A28131 PT-NANBH mr
2	1770	100.0	1770	6 A32184	A32184 NANBH PT p
3	1770	100.0	1770	6 ARI44032	ARI44032 Sequence
4	1763.6	99.6	3750	6 A28157	A28157 PT-NANBH mr
5	1763.6	99.6	3750	6 A32203	A32203 NANBH PT p
6	1763.6	99.6	3750	6 ARI44051	ARI44051 Sequence
7	1580.6	89.3	9359	14 ARI13916	ARI13916 Hepatitis C
8	1570.6	88.7	9373	14 ARI208024	ARI208024 Hepatitis C
9	1562.6	88.3	9357	14 ARI587845	ARI587845 Hepatitis C
10	1562.6	88.3	9410	14 ARI587845	ARI587845 Hepatitis C
11	1562.6	88.3	9547	14 ARI587845	ARI587845 Hepatitis C
12	1557.8	88.0	9374	14 ARI587845	ARI587845 Hepatitis C
13	1557	88.0	9425	14 ARI587845	ARI587845 Hepatitis C
14	1557	88.0	9573	14 ARI587845	ARI587845 Hepatitis C
15	1556.2	87.9	9410	14 ARI587845	ARI587845 Hepatitis C
16	1554.6	87.8	9379	14 ARI587845	ARI587845 Hepatitis C
17	1553.8	87.8	8638	6 ARI488105	ARI488105 Sequence
18	1553.8	87.8	8638	6 ARI488122	ARI488122 Sequence
19	1553.8	87.8	8638	6 ARI472297	ARI472297 Sequence

20	1553.8	87.8	8638	6 AX472315	AX472315 Sequence
21	1553.4	87.8	8642	6 ARI488101	ARI488101 Sequence
22	1553.4	87.8	8642	6 AX472292	AX472292 Sequence
23	1552.2	87.7	7989	6 ARI406043	ARI406043 Sequence
24	1552.2	87.7	7989	6 AX036255	AX036255 Sequence
25	1552.2	87.7	7989	12 SSE242652	AJ242652 Hepatitis C
26	1552.2	87.7	8001	6 ARI406042	ARI406042 Sequence
27	1552.2	87.7	8001	6 AX036254	AX036254 Sequence
28	1552.2	87.7	8001	12 SSE242654	AJ242654 Hepatitis C
29	1552.2	87.7	8637	6 ARI406041	ARI406041 Sequence
30	1552.2	87.7	8637	6 AX036253	AX036253 Sequence
31	1552.2	87.7	8637	12 SSE242651	AJ242651 Hepatitis C
32	1552.2	87.7	8638	6 ARI488104	ARI488104 Sequence
33	1552.2	87.7	8638	6 ARI488121	ARI488121 Sequence
34	1552.2	87.7	8638	6 AX472296	AX472296 Sequence
35	1552.2	87.7	8638	6 AX472314	AX472314 Sequence
36	1552.2	87.7	8639	6 ARI488100	ARI488100 Sequence
37	1552.2	87.7	8639	6 AX472291	AX472291 Sequence
38	1552.2	87.7	8643	6 ARI488102	ARI488102 Sequence
39	1552.2	87.7	8643	6 AX472294	AX472294 Sequence
40	1552.2	87.7	8649	6 ARI406044	ARI406044 Sequence
41	1552.2	87.7	8649	6 AX036256	AX036256 Sequence
42	1552.2	87.7	8649	12 SSE242653	AJ242653 Hepatitis C
43	1552.2	87.7	9033	14 HCT238800	AJ238800 Hepatitis C
44	1552.2	87.7	9410	14 HPCX182	D50485 Hepatitis C
45	1552.2	87.7	9605	6 AX739971	AX739971 Sequence

ALIGNMENTS

RESULT 1
A28131
LOCUS PT-NANBH mRNA fragment from patent GB2239245.
DEFINITION A28131
ACCESSION A28131
VERSION A28131.1 GI:1248616
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 1770)
AUTHORS
TITLES Post-translational non-A non-B hepatitis viral polypeptides
JOURNAL Patent: GB 2239245-A 3 26-JUN-1991;
FEATURES
source Location/Qualifiers
1..1770
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/clone="UG2"
<1..>1770
/catal_start=1
/catal_table=11
/product="PT-NANBH polypeptide"
/protein_id="CAA01925.1"
/db_xref="GI:1248617"
/translation="QNDPFDADLIEANILMRHEMGDITRVSEBNKVIILDSFDPIRA
EEDEREVSVAPEILKSKKFPAMPAMPADVPNPPLBEMKAPDYVPVHGGFLPPT
KTPPIPPRRKRTVITESTVSALAEATKAFSGSPAVSGTATAPDSDGCG
ACSDVESVSMPELGEEDPDPLSDSGSTVEEEDGVCMSGTATGAILTPCA
ESKUPINLNSILRHHNMVYATTSRSASQOKKVTPTRLQIIDLHYDVYKEMAK
ASTVRKLLSVBACKLTTPHSAKSFEFGADVRNLSSKALNHSISVEDLDEET
PIDITMANEVFCVQPERGKRPALIVFDLGVAVCRKMLYDVSTLPOAVGSS
YGVQVSQGVREFLVAMSKTTPMGFAVDFRCSFTVTENIRVEBSYOCCLAVE
ARQALISLTERLYIGPLTNSKGNGVRCRCAAGVLTTCNTLTCYKAAACRAA
KLODCTMIVCGGLIVTICGAGQEDPAASLRVFTAMTRYSPAPGPPPEVDLELIT
SCSNVSVADHDSGRVYTLTRDP"

ORIGIN
Query Match 100.0%; Score 1770; DB 6; length 1770;
Best local similarity 100.0%; Pred. No. 0;
Matches 1770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAATGACTTCCAGACGCTGACCTCATGAGGCCAACCCTCTGTGGCGGATGAGATG 60
|
|
|
Db 1 CAAATGACTTCCAGACGCTGACCTCATGAGGCCAACCCTCTGTGGCGGATGAGATG 60
|
|
|
QY 61 GGGGGGGAATTACCCGCGTGAAGTCAAGAACAGGTGTAATCTTGGACTTTTCGAC 120
|
|
|
Db 61 GGGGGGGAATTACCCGCGTGAAGTCAAGAACAGGTGTAATCTTGGACTTTTCGAC 120
|
|
|
QY 121 CCGCTCCAGCGGAGGATGAGGGGAAGTCCGTCGCGGAGATCCGCGGAA 180
|
|
|
Db 121 CCGCTCCAGCGGAGGATGAGGGGAAGTCCGTCGCGGAGATCCGCGGAA 180
|
|
|
QY 181 TCCAGAAATTCACACAGCGATGCCGATGGGCAAGCCGATTAACCTCCGCTG 240
|
|
|
Db 181 TCCAGAAATTCACACAGCGATGCCGATGGGCAAGCCGATTAACCTCCGCTG 240
|
|
|
QY 241 CTGAGACTCTGGAAGGCCCGGACTAGTCCCTCCAGTGTATCATGGTGCCCACTGCCA 300
|
|
|
Db 241 CTGAGACTCTGGAAGGCCCGGACTAGTCCCTCCAGTGTATCATGGTGCCCACTGCCA 300
|
|
|
QY 301 CCTACTAAGACCCCTCTATACACCTCCAGCGGAAGAGACAGTGTCTGACAGAA 360
|
|
|
Db 301 CCTACTAAGACCCCTCTATACACCTCCAGCGGAAGAGACAGTGTCTGACAGAA 360
|
|
|
QY 361 TCCACCGTGTCTTCTGCGCGAGCTTGCCACAAAGGCTTTTGATCTCCGACCG 420
|
|
|
Db 361 TCCACCGTGTCTTCTGCGCGAGCTTGCCACAAAGGCTTTTGATCTCCGACCG 420
|
|
|
QY 421 TGGCGCTGTGACAGCGGCAAGCGACCGCCCTCTGACCAATCTCTGCAAGCGCGGA 480
|
|
|
Db 421 TGGCGCTGTGACAGCGGCAAGCGACCGCCCTCTGACCAATCTCTGCAAGCGCGGA 480
|
|
|
QY 481 GCAGATCTGACGTTGATGTATTCCTCATGCCCCCTTGAGGGGAGCGCGGGAGAC 540
|
|
|
Db 481 GCAGATCTGACGTTGATGTATTCCTCATGCCCCCTTGAGGGGAGCGCGGGAGAC 540
|
|
|
QY 541 CCGGATCTGACGCGGCTTGCTGTACCGTGAAGTGAAGGCGCGTGAAGACGTCGTC 600
|
|
|
Db 541 CCGGATCTGACGCGGCTTGCTGTACCGTGAAGTGAAGGCGCGTGAAGACGTCGTC 600
|
|
|
QY 601 TGTGTGATGATCTCTACACATGAGAGGCGCTGTATCATCGCCATGCGCTGAGAGAA 660
|
|
|
Db 601 TGTGTGATGATCTCTACACATGAGAGGCGCTGTATCATCGCCATGCGCTGAGAGAA 660
|
|
|
QY 661 AGCAAGCTGCCATCAACGCGTTGAGCAACTCTTGTGTCGTCACCAACAATGCTTAC 720
|
|
|
Db 661 AGCAAGCTGCCATCAACGCGTTGAGCAACTCTTGTGTCGTCACCAACAATGCTTAC 720
|
|
|
QY 721 GCTACACATCCGACGCGCAAGCCAGCGGCAAGAGAGTCACTTTGACAGACTGCA 780
|
|
|
Db 721 GCTACACATCCGACGCGCAAGCCAGCGGCAAGAGAGTCACTTTGACAGACTGCA 780
|
|
|
QY 781 ATCTTGAAGATATCAACAGACGCTGTCAAGAGATGAAGGCCAGCGCTGCACAGTT 840
|
|
|
Db 781 ATCTTGAAGATATCAACAGACGCTGTCAAGAGATGAAGGCCAGCGCTGCACAGTT 840
|
|
|
QY 841 AAGGCTAAGCTTCTATCAGTGAAGAGCTGCAAGCTGACGCCCCCACTTCGCGCAA 900
|
|
|
Db 841 AAGGCTAAGCTTCTATCAGTGAAGAGCTGCAAGCTGACGCCCCCACTTCGCGCAA 900
|
|
|
QY 901 TCTAAATTTGGCTATGAGGCAAGAGCTCCGGAACCTATCCAGAGGCGCATTTAACAC 960
|
|
|
Db 901 TCTAAATTTGGCTATGAGGCAAGAGCTCCGGAACCTATCCAGAGGCGCATTTAACAC 960
|
|
|
QY 961 ATCCGCTCGGTGTGGAGAGCTTTGTGGAAGACCTGAAACCAATTTGACACCACTATC 1020
|
|
|
Db 961 ATCCGCTCGGTGTGGAGAGCTTTGTGGAAGACCTGAAACCAATTTGACACCACTATC 1020
|
|
|
QY 1021 ATGGCAAAATGAGGGTTTCTGCGTCAACAGAGAGAGGGCGCAAGCGAGCTGCG 1080
|
|
|
Db 1021 ATGGCAAAATGAGGGTTTCTGCGTCAACAGAGAGAGGGCGCAAGCGAGCTGCG 1080
|
|
|

QY 1081 CTTATGATGTTCCAGACTTGGGGGTCCGTGTGTGTCGAGAAATGCCCCCTATGACGCTG 1140
|
|
|
Db 1081 CTTATGATGTTCCAGACTTGGGGGTCCGTGTGTGTCGAGAAATGCCCCCTATGACGCTG 1140
|
|
|
QY 1141 GTCTCACCCCTCCCTCAGGCTGTGATGGGCTTCCTGTACGGATTCCAGTATTTCTCTGGA 1200
|
|
|
Db 1141 GTCTCACCCCTCCCTCAGGCTGTGATGGGCTTCCTGTACGGATTCCAGTATTTCTCTGGA 1200
|
|
|
QY 1201 CAGCGGATCGAGTTCTGTGTGAACGCTGGAATCAAGAAAGAACCCCTATGGGCTTTTGA 1260
|
|
|
Db 1201 CAGCGGATCGAGTTCTGTGTGAACGCTGGAATCAAGAAAGAACCCCTATGGGCTTTTGA 1260
|
|
|
QY 1261 TATGACACCCGCTGTTTGACTCAACAGTCACTGAAATGACATCCGTGTAAGAGTCA 1320
|
|
|
Db 1261 TATGACACCCGCTGTTTGACTCAACAGTCACTGAAATGACATCCGTGTAAGAGTCA 1320
|
|
|
QY 1321 ATTTATCAATGTTGTGACTTGGCCCCGGAAGCAGACAGGCCATTAAGTCCCTCAGAG 1380
|
|
|
Db 1321 ATTTATCAATGTTGTGACTTGGCCCCGGAAGCAGACAGGCCATTAAGTCCCTCAGAG 1380
|
|
|
QY 1381 CCGCTTATATCGGGGATCCCTGTACTAATCAAAAGGGAGAACTGGCGTATGCGCG 1440
|
|
|
Db 1381 CCGCTTATATCGGGGATCCCTGTACTAATCAAAAGGGAGAACTGGCGTATGCGCG 1440
|
|
|
QY 1441 TGCCGCGGAGCGCGGTCTGACGACTAGCTGCGGTAAATCCCTCAATGTTACTTGAAG 1500
|
|
|
Db 1441 TGCCGCGGAGCGCGGTCTGACGACTAGCTGCGGTAAATCCCTCAATGTTACTTGAAG 1500
|
|
|
QY 1501 GCTTGTGAGCTGTGTGAGCTGTCAAGAGTCCAGAGCTGCAAGATGCTGTGTGCGGAGAC 1560
|
|
|
Db 1501 GCTTGTGAGCTGTGTGAGCTGTCAAGAGTCCAGAGCTGCAAGATGCTGTGTGCGGAGAC 1560
|
|
|
QY 1561 GGCCTTGTGTTATCTGAGAGCGCGGGAACCCAGAGAGCGCGGCGACTACAGATC 1620
|
|
|
Db 1561 GGCCTTGTGTTATCTGAGAGCGCGGGAACCCAGAGAGCGCGGCGACTACAGATC 1620
|
|
|
QY 1621 TTACAGAGGCTATGACTAGGTACTTGCCCCCCCCGGGAGCCGCCCCAACAAGATAC 1680
|
|
|
Db 1621 TTACAGAGGCTATGACTAGGTACTTGCCCCCCCCGGGAGCCGCCCCAACAAGATAC 1680
|
|
|
QY 1681 GACCTGAGGTTGATTAACATATGCTCTCCCAATGTGTGCGGAGAGATGATCTGGC 1740
|
|
|
Db 1681 GACCTGAGGTTGATTAACATATGCTCTCCCAATGTGTGCGGAGAGATGATCTGGC 1740
|
|
|
QY 1741 AAAAGGTAATACTACCTCAACCGTGACCG 1770
|
|
|
Db 1741 AAAAGGTAATACTACCTCAACCGTGACCG 1770
|
|
|

RESULT 2
A32184 1770 bp RNA linear PAT 10-DEC-1996
LOCUS
DEFINITION NAMBHV PT polypotein fragment.
ACCESSION A32184
VERSION A32184.1 GI:1926538
KEYWORDS
SOURCE
ORGANISM Non-A, non-B hepatitis virus
Non-A, non-B hepatitis virus
Virus; unclassified viruses.
I (bases 1 to 1770)
REFERENCE
AUTHORS Patenc: FR 2655990-A 3 21-JUN-1991;
JOURNAL Location/Qualifiers
FEATURES
source
1..1770
/organism="Non-A, non-B hepatitis virus"
/mol_type="unassigned RNA"
/db_xref="taxon:12440"
<1..>1770
/codon_start=1
/product="partial NAMBHV-PT polypotein"
/protein_id="CAA02088.1"
/db_xref="GI:1926539"
/translation="QNDFPDADLLEANLWHEWKGDI TRVESKNKVILDSFPLRA

REFERENCE	UnClassified.				
AUTHORS	1 (bases 1 to 1770)				
TITLE	Hightfield, P. Edmund., Rodgers, B. Colin., Tedder, R. Seton. and				
JOURNAL	Barbara, J. Anthony, James.				
FEATURES	PT-NANB hepatitis polyprotein				
source	Patent: US 6210675-A 3 03-APR-2001;				
	Location/Qualifiers				
	1..1770				
ORIGIN	/organism="unknown"				
	/mol_type="unassigned DNA"				
Query Match	100.0%;	Score 1770;	DB 6;	Length 1770;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 1770;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1	CAAAATGACTTCCAGACGCTGACCTCATGAGGCCAACCTCTGTGGCGCATGAGATG	60		
DB	1	CAAAATGACTTCCAGACGCTGACCTCATGAGGCCAACCTCTGTGGCGCATGAGATG	60		
QY	61	GGGGGGGACATTACCGCGGTGAGTGAAGAACAAAGTAGTAATCTCTGACTTTCCAG	120		
DB	61	GGGGGGGACATTACCGCGGTGAGTGAAGAACAAAGTAGTAATCTCTGACTTTCCAG	120		
QY	121	CCGCTCCGAGCGGAGGAGTGAAGCGGAAAGTGTCCGTCCCGCGGAGATCTCTGCGAAA	180		
DB	121	CCGCTCCGAGCGGAGGAGTGAAGCGGAAAGTGTCCGTCCCGCGGAGATCTCTGCGAAA	180		
QY	181	TCCAGAAATTTCCACAGCGATGCCCGATGGGCAAGCCCGGATTTACACCTCCGCTG	240		
DB	181	TCCAGAAATTTCCACAGCGATGCCCGATGGGCAAGCCCGGATTTACACCTCCGCTG	240		
QY	241	CTGAGATCTCGAAGGCGCCGAGCTAGTCCCTCCAGGTGATCATGGGGCCACAGCGCA	300		
DB	241	CTGAGATCTCGAAGGCGCCGAGCTAGTCCCTCCAGGTGATCATGGGGCCACAGCGCA	300		
QY	301	CCTACTAAGACCCCTCTATACCACTTCAACGAGAAAGAGACAGTTGTCTGACAGAA	360		
DB	301	CCTACTAAGACCCCTCTATACCACTTCAACGAGAAAGAGACAGTTGTCTGACAGAA	360		
QY	361	TCCACCGTGTCTTTCGCTGCGGAGCTTTCGCAAAAGCTTTTGTGCTCCGACCG	420		
DB	361	TCCACCGTGTCTTTCGCTGCGGAGCTTTCGCAAAAGCTTTTGTGCTCCGACCG	420		
QY	421	TGCGCGGTGACAGCGGCAAGCGGCAAGCGGCGCTCTGACCAATCTCCAGCGCGGCA	480		
DB	421	TGCGCGGTGACAGCGGCAAGCGGCAAGCGGCGCTCTGACCAATCTCCAGCGCGGCA	480		
QY	481	GCAGATCTGACGTTGAGTGTATTCCTCATGCCCCCTTGAAGGGAGCGCGGGAGC	540		
DB	481	GCAGATCTGACGTTGAGTGTATTCCTCATGCCCCCTTGAAGGGAGCGCGGGAGC	540		
QY	541	CCCGATCTCAGCGAGCGGCTTGTGTCTACCGTGAAGAGCGCGGTGAAGACGTCGTC	600		
DB	541	CCCGATCTCAGCGAGCGGCTTGTGTCTACCGTGAAGAGCGCGGTGAAGACGTCGTC	600		
QY	601	TGCTGTGATGTCTTACATGAGAGGCGCTCTATACAGCCATGCGCTGCGAGAGAA	660		
DB	601	TGCTGTGATGTCTTACATGAGAGGCGCTCTATACAGCCATGCGCTGCGAGAGAA	660		
QY	661	AGCAAGCTGCCATCAACGCGTTGAGCAACTTTTGTCTGCTGATCCACAAACATGCTAC	720		
DB	661	AGCAAGCTGCCATCAACGCGTTGAGCAACTTTTGTCTGCTGATCCACAAACATGCTAC	720		
QY	721	GCTACCAATCCCGAGCGGCAAGCGGCGGAGAAAGTCACTTTTGAAGAATGCGAA	780		
DB	721	GCTACCAATCCCGAGCGGCAAGCGGCGGAGAAAGTCACTTTTGAAGAATGCGAA	780		
QY	781	ATCCTGAGAGTCACTACAGAGAGCTGTCAAGAGATGAGAGCGAGGCGTCCACAGTT	840		
DB	781	ATCCTGAGAGTCACTACAGAGAGCTGTCAAGAGATGAGAGCGAGGCGTCCACAGTT	840		
QY	841	AAGGCTAAGCTTCTATCAGTAGAGAAAGCGCTGCAAGCTGACGCCCCACATTCGCGCAA	900		

DB	841	AAGGCTAAGCTTCTATCAGTAGAGAGAACCTGCAAGCTGAGCGCCCACTTCGCGCAA	900		
QY	901	TCTAAATTTGGCTATAGGGCAAGAGCTCCGAAACCTATCCAGACAGCCATTACCAAC	960		
DB	901	TCTAAATTTGGCTATAGGGCAAGAGCTCCGAAACCTATCCAGACAGCCATTACCAAC	960		
QY	961	ATCCGCTCGTGTGGAGAGACTTGTGGAAGACCTGAACACCAATTGACACGATC	1020		
DB	961	ATCCGCTCGTGTGGAGAGACTTGTGGAAGACCTGAACACCAATTGACACGATC	1020		
QY	1021	ATGCAAAAAATGAGGTTTTCTGCGTCCAGACAGAGAGAGAGCGCGCAAGCACTCGC	1080		
DB	1021	ATGCAAAAAATGAGGTTTTCTGCGTCCAGACAGAGAGAGAGCGCGCAAGCACTCGC	1080		
QY	1081	CTTATCGTGTCCAGACTTGGGGGTCCGTGTGTGCGGAAATGAGCCCTCTATGACGTG	1140		
DB	1081	CTTATCGTGTCCAGACTTGGGGGTCCGTGTGTGCGGAAATGAGCCCTCTATGACGTG	1140		
QY	1141	GTCTCACCCCTCCAGGCTGTGATGGGCTCCTGTACGGAATTCAGTATTCCTGGA	1200		
DB	1141	GTCTCACCCCTCCAGGCTGTGATGGGCTCCTGTACGGAATTCAGTATTCCTGGA	1200		
QY	1201	CAGCGGTCGAGTTCCTGTGAAAGCTGGAATCAAGAAAGAGAGCCCTATGGCTTTGCA	1260		
DB	1201	CAGCGGTCGAGTTCCTGTGAAAGCTGGAATCAAGAAAGAGAGCCCTATGGCTTTGCA	1260		
QY	1261	TATGACACCCGCTGTTTGAATCAAGTCACTGAGATGACATCCGTGTAGAGAGTCA	1320		
DB	1261	TATGACACCCGCTGTTTGAATCAAGTCACTGAGATGACATCCGTGTAGAGAGTCA	1320		
QY	1321	ATTATCAATGTGATGCTTGGGCGCCGAGACAGAGAGGCAATGAGTCCGTCAGAG	1380		
DB	1321	ATTATCAATGTGATGCTTGGGCGCCGAGACAGAGAGGCAATGAGTCCGTCAGAG	1380		
QY	1381	CGGCTTTATTCGGGGGTCCCTGACTAATTCAAAGGGCAAGCTGCGCTATGCGCG	1440		
DB	1381	CGGCTTTATTCGGGGGTCCCTGACTAATTCAAAGGGCAAGCTGCGCTATGCGCG	1440		
QY	1441	TGCGCGGAGCGGCTGTGACGACTAGCTAGCTGCGGTATATCCCTCAATGTTACTTGAAG	1500		
DB	1441	TGCGCGGAGCGGCTGTGACGACTAGCTAGCTGCGGTATATCCCTCAATGTTACTTGAAG	1500		
QY	1501	GCTCTGAGAGCTGTGAGAGCTGCAAGAGCTCAAGAGCTCAAGATGCTGTGCGGAGAC	1560		
DB	1501	GCTCTGAGAGCTGTGAGAGCTGCAAGAGCTCAAGAGCTCAAGATGCTGTGCGGAGAC	1560		
QY	1561	GGCTTGTCTGATCTGAGAGCGGAGAACCCAGAGAGAGCGCGGAGCTTACGATGTC	1620		
DB	1561	GGCTTGTCTGATCTGAGAGCGGAGAACCCAGAGAGAGCGCGGAGCTTACGATGTC	1620		
QY	1621	TTCAAGAGAGCTATGACATGATCTTTCGCCCCCGGGAGACCGGCCCAACCAAGATAC	1680		
DB	1621	TTCAAGAGAGCTATGACATGATCTTTCGCCCCCGGGAGACCGGCCCAACCAAGATAC	1680		
QY	1681	GACCTGAGATGATTAATCAATGCTCTCAATGATGTGCGGCGAGATGATCTTGGC	1740		
DB	1681	GACCTGAGATGATTAATCAATGCTCTCAATGATGTGCGGCGAGATGATCTTGGC	1740		
QY	1741	AAAAAGGTATTAATCAATCACTCAACCGGTGACCG	1770		
DB	1741	AAAAAGGTATTAATCAATCACTCAACCGGTGACCG	1770		
RESULT 4					
A28157					
LOCUS	A28157 3750 bp DNA linear PAT 04-AUG-1995				
DEFINITION	PT-NANBH mRNA fragment from patent GB229245.				
ACCESSION	A28157				
VERSION	A28157.1 GI:1248640				
KEYWORDS					
SOURCE	synthetic construct				
ORGANISM	synthetic construct				

other sequences; artificial sequences.

1 (bases 1 to 3750)

REFERENCE

Post-translational non-A non-B hepatitis viral polypeptides

Patent: GB 2239245-A 22 JUN-1991

Location/Qualifiers

FEATURES

1..3750

/organism="synthetic construct"

/mol_type="unassigned DNA"

/db_xref="taxon:32630"

/note="contig formed by cDNA clones from 3' end"

CDS

<1..3750

/codon_start=1

/transl_table=1

/product="PT-NANB polypeptide"

/protein_id="CA01932.1"

/db_xref="GI:1248641"

/translation="MEGVFTGLTHVDHFLSOTKQAGDNPEYLVAYQATVCARQAP
PSMDWMKCLIRKPTLRGPTPLRYLGAQNEVTLTHPTTKTMAQMSADLEVTST
MYLVGVLAALAYCLTGSVTVGRILISGRPAIYDPEVLYQEDMEQASHLPY
IEGMQLAQFKOKALGLQATKQABAAPVESKRALETVMACHMNFISGTYL
AGLSTLPGNPALASLMAFTASVTPLTOSTLNLILGWAQALAPPSAASFVAG
IAGAAYGSLGKVLVDILAGVAGVAGALVAFKVMGSEMPSTEDLVNLLPALISFA
LVGVCAALIRHVGPEGAVOMNRLIAFSGNHVSPTHYVPESDAARVOTLS
DLTITQLKRLHQMINDGSTPCSGMLRDMVDMICTVLADFTWLQSLPLPQVP
FESQGRYKWRGDCIMOTTCSCQAQITGHVNGSMRIYEPKCSMMHRTPIINAY
TTGPTCPSPAPNTSRALMRYAAEYEVTRKVDGFHYVTSMTDVAQPCQVPAPEFT
EVDGVLHRYAPACKPLREVTFOVGLVGSQLPCEPEPDAVLSMLTDPHI
TAEATKRLARGSPSLASSASQASPSKATYITONFPDPLLEAMLMHENGQ
DIATRESENKVLILDSFDPLRAEDEREVSYPALIRKSKPPAMPAMRPNYPL
LESWKAQVVPVYHGCPLPTKTPPIPERARKVTLRESYTSALALATAPGSS
EBSAVDSGRITAPDPDPSDGGAGSVESISNRPLEGEQDPLSDGSSTYSBAG
EDVACCSMSYTMGALITPCAESKLPILNLSNLRHNMVATTSBSAQKQKV
TEPRQLIDHDYQDVLKEMKAKATVAKLISVEACKLTPPSASKSGYKQKVN
LSKKAINHRSVWEDLEDETETPIDTTIAKNEVFCVQERGRKAPALVFDLCVR
VCEKMAIDVYSTLPQAVMSSYGFQSPQRYEPLVNMVSKSKTGMGFAYDRGDS
TYTENDIRVESIYQCCDLAPEAROALRSITRLVYGGPTLVLTSGNKGQRCRARSAG
LTTSCNTLTCYIKASACRAALQDCTMIVCGDGLVIVCESAGTEBDASLRVFTFA
MTRYSAPGDPPEYDELITISSSNSVVAHDASGRVYVLTDP"

ORIGIN

Query Match 99.64; Score 1763.6; DB 6; Length 3750;

Best Local Similarity 99.64; Pred. No. 0;

Matches 1766; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

1 CAAATGACTTCCAGACGCTGACTCATGAGGCCAACTCTGTGGCGGATGAGATG 60
1981 CAAATGACTTCCAGACGCTGACTCATGAGGCCAACTCTGTGGCGGATGAGATG 2040
61 GCGGGGACATTACCCGCGTGAAGTCAAGAGAAAGAGTAATCTCTGACTTTTCGAC 120
2041 GCGGGGACATTACCCGCGTGAAGTCAAGAGAAAGAGTAATCTCTGACTTTTCGAC 2100
121 CGGCTCGAGCGGAGGATGAGCGGGAAGTGCCTCCGGCGGAGATCTGCGGAAA 180
2101 CGGCTCGAGCGGAGGATGAGCGGGAAGTGCCTCCGGCGGAGATCTGCGGAAA 2160
181 TCCAGAAATTTCCACAGCGATGCGCGCATGGGCAAGCCGGATTAACAACCTCCGCTG 240
2161 TCCAGAAATTTCCACAGCGATGCGCGCATGGGCAAGCCGGATTAACAACCTCCGCTG 2220
241 CTGGAATCTCTGAAAGCCCGGACCTAGCTCTCTCCAGTGTATCATGGGTGCCATGCGCA 300
2221 CTGGAATCTCTGAAAGCCCGGACCTAGCTCTCTCCAGTGTATCATGGGTGCCATGCGCA 2280
301 CCTACTAAGACCCCTCTCTATACACCTCCAGGAGAAAGAGCAGTGTGTTGACAGAA 360
2281 CCTACTAAGACCCCTCTCTATACACCTCCAGGAGAAAGAGCAGTGTGTTGACAGAA 2340
361 TCCACGCTGCTTCTGCGCTGAGGAGCTTGCAAAAGGCTTTGGTAGCTCCGAGACG 420
2341 TCCACGCTGCTTCTGCGCTGAGGAGCTTGCAAAAGGCTTTGGTAGCTCCGAGACG 2400

421 TCGCCGCTGACAGCGGCAAGGCAAGCCGCCCTCTGACCAATTCCTCGACGAGCGGCA 480
2401 TCGCCGCTGACAGCGGCAAGGCAAGCCGCCCTCTGACCAATTCCTCGACGAGCGGCA 2460
481 GCAGATCTGACGCTTGAAGTGTATTCCTCATGCCCCCTTGAAGGAGAGCCGGGAGAC 540
2461 GCAGATCTGACGCTTGAAGTGTATTCCTCATGCCCCCTTGAAGGAGAGCCGGGAGAC 2520
541 CCCGATCTGACGAGCGGCTTGTGTATCCGTGAGTGAAGAGCCGCTGAGACCTGCTGC 600
2521 CCCGATCTGACGAGCGGCTTGTGTATCCGTGAGTGAAGAGCCGCTGAGACCTGCTGC 2580
601 TGCTGCTCATGTCCTACACATGAGCAGGCGCTGTATACGCGCATAGCGCTGCGAGGAA 660
2581 TGCTGCTCATGTCCTACACATGAGCAGGCGCTGTATACGCGCATAGCGCTGCGAGGAA 2640
661 AGCAAGCTGCCATCAACCGCTTGAACCACTCTTGTGCTGCTGACCAACAATGCTCTAC 720
2641 AGCAAGCTGCCATCAACCGCTTGAACCACTCTTGTGCTGCTGACCAACAATGCTCTAC 2700
721 GCTACCAATCCCGCAGCGCAAGCCGCGGCGAGAGAGTCACTTTGACAGACTGCA 780
2701 GCTACCAATCCCGCAGCGCAAGCCGCGGCGAGAGAGTCACTTTGACAGACTGCA 2760
781 ATCTGAGAGATCCTACACAGGACGTCCTCAAGAGATGAAGGCGAGGCGTCCACAT 840
2761 ATCTGAGAGATCCTACACAGGACGTCCTCAAGAGATGAAGGCGAGGCGTCCACAT 2820
841 AAGGCTAAGCTTCTATCACTAGAGAGAGCTCAAGCTGACCGCCCACTTTCGAGCA 900
2821 AAGGCTAAGCTTCTATCACTAGAGAGAGCTCAAGCTGACCGCCCACTTTCGAGCA 2880
901 TCTAAATTTGGCTATGAGGCAAGAGCGTCCGAACTATCCAGAGGCTATTAAACAC 960
2881 TCTAAATTTGGCTATGAGGCAAGAGCGTCCGAACTATCCAGAGGCTATTAAACAC 2940
961 ATCCGCTCGGTGAGGAGACTTGTGGAAGACCTGAACACCAATTGACACCAACATC 1020
2941 ATCCGCTCGGTGAGGAGACTTGTGGAAGACCTGAACACCAATTGACACCAACATC 3000
1021 ATGGCAAAAATGAGGCTTTCTGCGTCAACAGAGAGAGAGCGGCAAGCAGCTCCG 1080
3001 ATGGCAAAAATGAGGCTTTCTGCGTCAACAGAGAGAGAGCGGCAAGCAGCTCCG 3060
1081 CTATGCTGTTCCAGACTTGGGAGTCCGTGTGTGAGAGAAATGCGCTTATGACGTG 1140
3061 CTATGCTGTTCCAGACTTGGGAGTCCGTGTGTGAGAGAAATGCGCTTATGACGTG 3120
1141 GTCCTCAACCTCCCTCAGGCTGTGATGGGCTCCTGTAGGATTCAGATTCCTCTGGA 1200
3121 GTCCTCAACCTCCCTCAGGCTGTGATGGGCTCCTGTAGGATTCAGATTCCTCTGGA 3180
1201 CAGCGGCTGAGATCTCTGCTGAGACGCTGGAATCAAGAAACCCCTATGAGCTTTGCA 1260
3181 CAGCGGCTGAGATCTCTGCTGAGACGCTGGAATCAAGAAACCCCTATGAGCTTTGCA 3240
1261 TATGACACCGCTGTTTGAATCAACAGTCACTGAGATGACATCCGTGTAGAGAGTCA 1320
3241 TATGACACCGCTGTTTGAATCAACAGTCACTGAGATGACATCCGTGTAGAGAGTCA 3300
1321 ATTATCAATGTTTGAATCTTGGCCCCCGGAGGCAAGCCATTAAGTGTGCTCACAGAG 1380
3301 ATTATCAATGTTTGAATCTTGGCCCCCGGAGGCAAGCCATTAAGTGTGCTCACAGAG 3360
1381 CGGCTTATATGAGGAGTCCCTGACTAATCAAAAAGGCAAGCTGCGGCTATGCGCGG 1440
3361 CGGCTTATATGAGGAGTCCCTGACTAATCAAAAAGGCAAGCTGCGGCTATGCGCGG 3420
1441 TGCCTGCGAGCGGCTGCTGACAGTACGCTCGGTATATCCCTCAATGTTTACTTGAAG 1500
3421 TGCCTGCGAGCGGCTGCTGACAGTACGCTCGGTATATCCCTCAATGTTTACTTGAAG 3480
1501 GCCTCTGACGCTGTGTGAGCTGCAAAAGCTCCAGAGACTGACGATGTGTGTGCGGAGAC 1560

Db 3481 GCCTGACGCGCTGAGCTGCAAGCTGCAAGACTGTCAGTCTGTCGCGAGAC 3540
Qy 1561 GGCCTGTGTTATCTGTAGAGCGCGGAAACCCAGAGAGCGCGCGAGCTTACAGATC 1620
Db 3541 GGCCTGTGTTATCTGTAGAGCGCGGAAACCCAGAGAGCGCGCGAGCTTACAGATC 3600
Qy 1621 TTCACGAGGCTATGACTGATGCTGCCCCCGCGGAGCCCGCCCAACAGAAATAC 1680
Db 3601 TTCAGGAGGCTATGACTGATGCTGCCCCCGCGGAGCCCGCCCAACAGAAATAC 3660
Qy 1681 GACCTGAGTGTAAATCATGATGCTCTCCATGATGTCGTCGCGAGCATGATGTC 1740
Db 3661 GACCTGAGTGTAAATCATGATGCTCTCCATGATGTCGTCGCGAGCATGATGTC 3720
Qy 1741 AAAAGGGTATGACTGATGCTGCCCCCGCGGAGCCCGCCCAACAGAAATAC 1770
Db 3721 AAAAGGGTATGACTGATGCTGCCCCCGCGGAGCCCGCCCAACAGAAATAC 3750

RESULT 5
A32203
LOCUS A32203 3750 bp DNA linear PAT 10-DEC-1996
DEFINITION NANBH PT polypeptide fragment.
ACCESSION A32203
VERSION A32203.1 GI:1926564
KEYWORDS
SOURCE Non-A, non-B hepatitis virus
ORGANISM Non-A, non-B hepatitis virus
REFERENCE Viruses; unclassified viruses.
1 (bases 1 to 3750)
AUTHORS Patent: FR 2655990-A 22.21-JUN-1991;
JOURNAL Location/Qualifiers
FEATURES
source 1..3750
CDS
1..3750
/organism="Non-A, non-B hepatitis virus"
/mol_type="unassigned DNA"
/db_xref="taxon:12440"
<1..>3750
/codon_start=1
/product="partial NANBH-PT polypeptide"
/protein_id="CAA02095.1"
/db_xref="GI:1926565"
translation="MEGVFTGLTHVDHFLSQTOKADNFPYLAVOATVCAQAAP
PSMDMMKCLIRKLTGRLPTLYRLGVONETLPHITKFIACASADLEVYTS
WLVGVLAALAAVCLTGSVYIVGRILLSGPAIVREVYQSDMERGASLPL
IEGQMQLAEQKQXALGLIQTATKQABAAVYESKRLIEFKAKMNFISGTOYL
AGLSTLPGNPATIASLMAFTASVTSPLTQSTLLNLILGAVPNLSLPSAABFVAG
IAGAAGVSGIGAKVLDILAGYAGVAGLVAFKVMSGMBSTEDLVNLPALISGA
LVGVCAAILLRHYPGEGAVQMMNRLLAFASRGNHSPTHYVESDAAAVTQILS
DLITLQRLKHLINEDSTPGSWMRLVDWMICTVLADKTMLOSGLPRLGVP
FESCGRYGVNRGIMQOTSCGAOITGHVNGSMITVGRKTSNMHHGPPINAY
TTGPTCPSPAPYSRALNRVAEVEYTRVGDHYVTSMITDNKCPQYPALPFT
EVDGRLHRYAPACKPLREEVTFQVGLNQLVGSQLEPEPDAVAVITSLTDESHI
TAETAKRLRAGSPSLSSASQLSABSSKATYITQNDFPADILINLILRMHMG
DITRYSEKRVILDSFDPLRAEEDEREVSVAELIRKSKPPAMPAMPARDPYPL
LESMAADYPPVHGCPLEPTKTPPIPPRRKRVVLTESVSLALELTKARAGS
BSAAYDSGATAPDOPSPDGDGAGDVSYSMPLEGRPDLSDSKSVESBAG
EDVCCSMSTYTGALITPCABESKLPINALSNLSLHRHNVIVITRSKVSQKQV
TFDRLOILDDHYDVLKEMKAKASTVKALSVESACKLTPHSAKSKFEGAGKVRN
LSKAIINLRVSWEDLBDTEPDIITIMAKNEVCVOPBERGRAPALIYFDPILGR
VCEKALYDVVSTLPQAVGSSYGFQYSGQVFEFLVNAKSKTPMGFAYDRCFDS
TYTENDIRVESIYCCDLAPBAROAIISLTERLYIGGLTMSKGNCGVYRCRAGSV
LTTSCGNTLTCYLKASACRAKLODCMTLVCGDLIVYICSAAGTQEDBAASLRVTEA
MTRYSAPDEDPQPEYDLELTSCSSNSVVAHDAGSKRYIYITRDP"

ORIGIN

Query Match 99.6%; Score 1763.6; DB 6; Length 3750;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1766; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CAAATGACTTCCAGAGCGCTGACTCATGAGCGCAACCTCTGTCGCGAGATGATG 60
|||||
Db 1 CAAATGACTTCCAGAGCGCTGACTCATGAGCGCAACCTCTCTGTCGCGAGATG 60
|||||

Db 1981 CAAATGACTTCCAGAGCGCTGACTCATGAGCGCAACCTCTCTGTCGCGAGATG 2040
Qy 61 GCGCGGGAATTTACCCCGGTGAGTCAAGAAACAAAGTAAATCTGACCTTTGAC 120
Db 2041 GCGCGGGAATTTACCCCGGTGAGTCAAGAAACAAAGTAAATCTGACCTTTGAC 2100
Qy 121 CCGCTCCGAGGAGAGAGATGAGCGGAGATGTCCTCCGCGCGAGATCTCGGAAA 180
Db 2101 CCGCTCCGAGGAGAGAGATGAGCGGAGATGTCCTCCGCGCGAGATCTCGGAAA 2160
Qy 181 TCAGAAATTTCCACACAGCATGCTCCGATGAGCAACCCGAGATTACACCTTCGCTG 240
Db 2161 TCAGAAATTTCCACACAGCATGCTCCGATGAGCAACCCGAGATTACACCTTCGCTG 2220
Qy 241 CTGAGTCTGGAAGGCCCCGGAATGATGCTCCGATGAGTCAATGGTGGCCATGCA 300
Db 2221 CTGAGTCTGGAAGGCCCCGGAATGATGCTCCGATGAGTCAATGGTGGCCATGCA 2280
Qy 301 CTTACTAAGACCCCTCTATACACCTTCACAGAGAAAGAGAGAGATGTTCTGACAGAA 360
Db 2281 CTTACTAAGACCCCTCTATACACCTTCACAGAGAAAGAGAGAGATGTTCTGACAGAA 2340
Qy 361 TCACCGTGTCTTCTGCTGCGAGGCTTCCCAAAAGCTTTGTGTAAGTCCGAGACG 420
Db 2341 TCACCGTGTCTTCTGCTGCGAGGCTTCCCAAAAGCTTTGTGTAAGTCCGAGACG 2400
Qy 421 TCGCGCGTCAAGAGGAGCGCAACCGCCCTCTGACCAATCTCCAGAGAGCGGAG 480
Db 2401 TCGCGCGTCAAGAGGAGCGCAACCGCCCTCTGACCAATCTCCAGAGAGCGGAG 2460
Qy 481 GCAGATCTGACGTTGATGCTATTCCTCATGCCCCCTTGAGAGGAGCCCGGAGAC 540
Db 2461 GCAGATCTGACGTTGATGCTATTCCTCATGCCCCCTTGAGAGGAGCCCGGAGAC 2520
Qy 541 CCCGATCTCAGCAGCGGCTTTGGTCTACCGTGAAGAGAGCGCGTGAAGAGTCTC 600
Db 2521 CCCGATCTCAGCAGCGGCTTTGGTCTACCGTGAAGAGAGCGCGTGAAGAGTCTC 2580
Qy 601 TGCTGCTGATGCTCTACATGAGAGAGGCGCTGATCAAGCGCATGCGGAGAGAA 660
Db 2581 TGCTGCTGATGCTCTACATGAGAGAGGCGCTGATCAAGCGCATGCGGAGAGAA 2640
Qy 661 AGCAAGCTGCCATCAACGCGTGGAGCACTTTTGTGCTGCTCAACCAATGGTCTAC 720
Db 2641 AGCAAGCTGCCATCAACGCGTGGAGCACTTTTGTGCTGCTCAACCAATGGTCTAC 2700
Qy 721 GCTACCATATCCGAGCGGACGAGCGGAGAGAGAGAGTCACTTTGACAGCTGCA 780
Db 2701 GCTACCATATCCGAGCGGACGAGCGGAGAGAGAGTCACTTTGACAGCTGCA 2760
Qy 781 ATCCGAGCGATCACTACAGAGAGTGTCAAGAGAGTGAAGGAGGCGTCCACAGTT 840
Db 2761 ATCCGAGCGATCACTACAGAGAGTGTCAAGAGAGTGAAGGAGGCGTCCACAGTT 2820
Qy 841 AAGGCTAAGCTTCTATCAGTGAAGAGAGCTGCAAGCTGACGCCCCCATTTCCGCGCAA 900
Db 2821 AAGGCTAAGCTTCTATCAGTGAAGAGAGCTGCAAGCTGACGCCCCCATTTCCGCGCAA 2880
Qy 901 TCTAATTTGGCTATGAGGAGCAAGAGAGCTCCGAACTTATCAGCAAGAGCCATTACAC 960
Db 2881 TCTAATTTGGCTATGAGGAGCAAGAGAGCTCCGAACTTATCAGCAAGAGCCATTACAC 2940
Qy 961 ATCCGCTCCGAGGAGAGAGCTTGTGGAACAACCTGAACCAATTGACACACATC 1020
Db 2941 ATCCGCTCCGAGGAGAGAGCTTGTGGAACAACCTGAACCAATTGACACACATC 3000
Qy 1021 ATGCAAAAAATGAGGTTTCTGCTCCACACAGAGAGAGGCGCGCAACGACTTCG 1080
Db 3001 ATGCAAAAAATGAGGTTTCTGCTCCACACAGAGAGAGGCGCGCAACGACTTCG 3060
Qy 1081 CTTATCGTTCCTCAGACTTGGGAGTCCGTGTGTCGAGAAATAGCCCTTATGACGTG 1140
Db 3061 CTTATCGTTCCTCAGACTTGGGAGTCCGTGTGTCGAGAAATAGCCCTTATGACGTG 3120

QY 1141 GTTCCACCCCTGCTGAGGTGATGGGCTCTGTAAGGATTCAGATTCTCTGGA 1200
 Db 3121 GTTCCACCCCTGCTGAGGTGATGGGCTCTGTAAGGATTCAGATTCTCTGGA 3180
 QY 1201 CAGCGGGTGAAGTCTCTGAGTGAAGCGCTGGAATCAAGAGAGACCCCTATGGGCTTTGGA 1260
 Db 3181 CAGCGGGTGAAGTCTCTGAGTGAAGCGCTGGAATCAAGAGAGACCCCTATGGGCTTTGGA 3240
 QY 1261 TATGACACCCGCTGTTTGACTCAACAGTCACTGAGATGACATCCGTGAGAGAGTCA 1320
 Db 3241 TATGACACCCGCTGTTTGACTCAACAGTCACTGAGATGACATCCGTGAGAGAGTCA 3300
 QY 1321 ATTATCAATGTTGATCTTGAGCCCGGAGACGAGCCATAAGTGTGCTCAAGAG 1380
 Db 3301 ATTATCAATGTTGATCTTGAGCCCGGAGACGAGCCATAAGTGTGCTCAAGAG 3360
 QY 1381 CGGCTTTATATGGGGGGTCCCTGACTAATTCAAAAGGAGAACTGGGCTATCCCGG 1440
 Db 3361 CGGCTTTATATGGGGGGTCCCTGACTAATTCAAAAGGAGAACTGGGCTATCCCGG 3420
 QY 1441 TGGCGCGCAGCGGCGCTGCTGACGACTAGCTCGGTATATCCTCATATGTTACTTGAAG 1500
 Db 3421 TGGCGCGCAGCGGCGCTGCTGACGACTAGCTCGGTATATCCTCATATGTTACTTGAAG 3480
 QY 1501 GGCTGCGACGCGTGTGAGCTGCAAGCTCCAGAGACTGACAGATGCTCGTGTGCGGAGAC 1560
 Db 3481 GGCTGCGACGCGTGTGAGCTGCAAGCTCCAGAGACTGACAGATGCTCGTGTGCGGAGAC 3540
 QY 1561 GGGCTTGTGCTTATCTGTGAGAGCGCGGAAACCCAGAGAGACGCGGAGCTTACGAGTAC 1620
 Db 3541 GGGCTTGTGCTTATCTGTGAGAGCGCGGAAACCCAGAGAGACGCGGAGCTTACGAGTAC 3600
 QY 1621 TTTCAGGAGGCTATGACTAGTACTCTGCCCCCCCCCGGAGACCCGCCCCAACCAGAAATAC 1680
 Db 3601 TTTCAGGAGGCTATGACTAGTACTCTGCCCCCCCCCGGAGACCCGCCCCAACCAGAAATAC 3660
 QY 1681 GACCTGGAATTTGATATCATCATCATGCTCTCCATATGTTGCGTGGCGGACAGATGATCTGGC 1740
 Db 3661 GACCTGGAATTTGATATCATCATCATGCTCTCCATATGTTGCGTGGCGGACAGATGATCTGGC 3720
 QY 1741 AAAAGGATTAATCACTCAACCCGTCGACCCG 1770
 Db 3721 AAAAGGATTAATCACTCAACCCGTCGACCCG 3750

RESULT 6
 ARI44051 3750 bp DNA linear PAT 08-AUG-2001
 LOCUS ARI44051 Sequence 22 from patent US 6210675.
 DEFINITION ARI44051
 ACCESSION ARI44051 GI:15105918
 VERSION ARI44051.1
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1. (bases 1 to 3750)
 AUTHORS Highfield, P. Edmund., Rodgers, B. Colin., Tedder, R. Seton. and
 TITLE PT-NANB hepatitis polyprotein
 JOURNAL Patent: US 6210675-A 22 03-APR-2001;
 FEATURES Location/Qualifiers
 1..3750
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN

Query Match 99.6%; Score 1763.6; DB 6; Length 3750;
 Best Local Similarity 99.8%; Freq. No. 0;
 Matches 1766; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 CAAAATGACTTCCAGACGCTGACCTCATGAGGCCAACCTCTGTGGCGGATGAGATG 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 1981 CAAAATGACTTCCAGACGCTGACCTCATGAGGCCAACCTCTGTGGCGGATGAGATG 2040
 QY 61 GGGCGGAGATTAACCCGCTGAGATCAGAGAACAAAGTATGTAATCTTGACCTTTGAC 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2041 GGGCGGAGATTAACCCGCTGAGATCAGAGAACAAAGTATGTAATCTTGACCTTTGAC 2100
 QY 121 CCGCTCCGAGCGAGAGATGAGCGGAGATGTCGCTCCGCGGAGATCTCTGCGGAA 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2101 CCGCTCCGAGCGAGAGATGAGCGGAGATGTCGCTCCGCGGAGATCTCTGCGGAA 2160
 QY 181 TCCAGAAATTTCCACACAGCGATGCCCGCATGGGACGCGCGGATTAACAACCTCGCTG 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2161 TCCAGAAATTTCCACACAGCGATGCCCGCATGGGACGCGCGGATTAACAACCTCGCTG 2220
 QY 241 CTGAGATCTGGAAGGCCCGGAGATACGTCTCTCAATGATGATGAGGTGCCCCATGCGCA 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2221 CTGAGATCTGGAAGGCCCGGAGATACGTCTCTCAATGATGATGAGGTGCCCCATGCGCA 2280
 QY 301 CCTACTAAGACCCCTCTTATACACCTCCACGAGAGAGAGAGAGATTTGTTGACAGAA 360
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2281 CCTACTAAGACCCCTCTTATACACCTCCACGAGAGAGAGAGAGATTTGTTGACAGAA 2340
 QY 361 TCACCGGTCTTCTGCGCTGCGGAGCTTGCACAAAGGCTTTGTTAGTCTCGGACCG 420
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2341 TCACCGGTCTTCTGCGCTGCGGAGCTTGCACAAAGGCTTTGTTAGTCTCGGACCG 2400
 QY 421 TGGCGCTGTGACAGCGGACCGGCAACCGGCTCTCTGACCAATCTCTCGACGACGCGGA 480
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2401 TGGCGCTGTGACAGCGGACCGGCAACCGGCTCTCTGACCAATCTCTCGACGACGCGGA 2460
 QY 481 GGAGATCTGAGCTTGAAGTGTATTCCTCATGAGCCGCTTGAAGGAGGAGCGGAGGAC 540
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2461 GGAGATCTGAGCTTGAAGTGTATTCCTCATGAGCCGCTTGAAGGAGGAGCGGAGGAC 2520
 QY 541 CCCGATCTTCAGCGAGCGGTCTTGTGTCTACCGTGAATGAGAGCGGATGAGACGTCGTC 600
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2521 CCCGATCTTCAGCGAGCGGTCTTGTGTCTACCGTGAATGAGAGCGGATGAGACGTCGTC 2580
 QY 601 TGTGCTGATGTCTCTACATGAGACAGGCGCTCTGATACGCCCATGCGCTGCGAGAA 660
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2581 TGTGCTGATGTCTCTACATGAGACAGGCGCTCTGATACGCCCATGCGCTGCGAGAA 2640
 QY 661 AGCAAGCTGCGCATCAAGCGCTTGAAGCACTTTGCTGCGTCAACCAATGGTCTAC 720
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2641 AGCAAGCTGCGCATCAAGCGCTTGAAGCACTTTGCTGCGTCAACCAATGGTCTAC 2700
 QY 721 GCTACACATCCCGAGCGCAAGCCAGCGGCGAGAAAGTCACTTTGACAGACTGCA 780
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2701 GCTACACATCCCGAGCGCAAGCCAGCGGCGAGAAAGTCACTTTGACAGACTGCA 2760
 QY 781 ATCTGAGCATCACTACAGAGACGTCTCAAGAGATGAGAGCGAGGCGTCCACAGTT 840
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2761 ATCTGAGCATCACTACAGAGACGTCTCAAGAGATGAGAGCGAGGCGTCCACAGTT 2820
 QY 841 AAGGCTAAGCTCTATCACTAGAGAGAGCGTCAAGCTGAGCGCCCAATTCGCGCA 900
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2821 AAGGCTAAGCTCTATCACTAGAGAGAGCGTCAAGCTGAGCGCCCAATTCGCGCA 2880
 QY 901 TCTAAATTTGGCTATGCGGCAAGAGCGTCCGGAACCTATCCAGAGAGCGCATTAACAC 960
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2881 TCTAAATTTGGCTATGCGGCAAGAGCGTCCGGAACCTATCCAGAGAGCGCATTAACAC 2940
 QY 961 ATCCGCTCCGTGTGAGAGGACTTTGTTGAAGACACTGAACCAATTGACACACCATC 1020
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2941 ATCCGCTCCGTGTGAGAGGACTTTGTTGAAGACACTGAACCAATTGACACACCATC 3000
 QY 1021 ATGGCAAAAATGAGGTTTTCGCGTCAACAGAGAGAGAGGCGCAAGCAGCTGCG 1080
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 3001 ATGGCAAAAATGAGGTTTTCGCGTCAACAGAGAGAGAGGCGCAAGCAGCTGCG 3060
 QY 1081 CTTATCGTGTCCAGACTTGGGAGGTCCGTGTGTCGAGAAATAGCCCTCTATGACGTG 1140
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 3061 CTTATCGTGTCCAGACTTGGGAGGTCCGTGTGTCGAGAAATAGCCCTCTATGACGTG 3120

```

Qy 1141 GTCCTCACCCCTCCCTCAGGCTGTGATGGGCTCCTCGTACGATTCAGATTTCTCCTGGA 1200
    |||
Db 3121 GCTCCACCCCTCCCTCAGGCTGTGATGGGCTCCTCGTACGATTCAGATTTCTCCTGGA 3180
Qy 1201 CAGCGGCTGTGATTTCTGTGTAAACGCTTGAAATCAAAGAAGACCCCTATGCGCTTTGCA 1260
    |||
Db 3181 CAGCGGCTGTGATTTCTGTGTAAACGCTTGAAATCAAAGAAGACCCCTATGCGCTTTGCA 3240
Qy 1261 TATGACACCCGCTGTTTGAATCAACAGTCACTGAGAAATGATCCGTGTAGAGAGTCA 1320
    |||
Db 3241 TATGACACCCGCTGTTTGAATCAACAGTCACTGAGAAATGATCCGTGTAGAGAGTCA 3300
Qy 1321 ATTATCAATGTTGTGACTTTGGCCCCCGAAGCAGACAGCCATTAAGTGTGCTACAGAG 1380
    |||
Db 3301 ATTATCAATGTTGTGACTTTGGCCCCCGAAGCAGACAGCCATTAAGTGTGCTACAGAG 3360
Qy 1381 CGGCTTTATATGCGGGGCTCCCTGACTAATTCAAAAGGCGAGACTGCGGCTATGCGCGG 1440
    |||
Db 3361 CGGCTTTATATGCGGGGCTCCCTGACTAATTCAAAAGGCGAGACTGCGGCTATGCGCGG 3420
Qy 1441 TGCCTGCGAGGCGGCGGTGCTGAGAGTACTGCGGTAATACCTCAATGTTACTTTGAAG 1500
    |||
Db 3421 TGCCTGCGAGGCGGCGGTGCTGAGAGTACTGCGGTAATACCTCAATGTTACTTTGAAG 3480
Qy 1501 GCCTCTGACGCTGTGAGCTGCAAGAGCTCAGAGACTGACAGATGCTGTGCGGAGAGC 1560
    |||
Db 3481 GCCTCTGACGCTGTGAGCTGCAAGAGCTCAGAGACTGACAGATGCTGTGCGGAGAGC 3540
Qy 1561 GGCCTTTGCTGTTATCTGTGAGAGCGCGGAGAACCCAGAGAGAGCGCGGCGACTTACGAGT 1620
    |||
Db 3541 GGCCTTTGCTGTTATCTGTGAGAGCGCGGAGAACCCAGAGAGAGCGCGGCGACTTACGAGT 3600
Qy 1621 TTCAGGAGGCTATGACTGACTGCTGCCCCCGCGGAGACCCGCCCAACAGAAATAC 1680
    |||
Db 3601 TTCAGGAGGCTATGACTGACTGCTGCCCCCGCGGAGACCCGCCCAACAGAAATAC 3660
Qy 1681 GACCTGAGATTGATTAACATCATGCTCTTCCAAATGTCGTCGCGCAGCATGATCTGGC 1740
    |||
Db 3661 GACCTGAGATTGATTAACATCATGCTCTTCCAAATGTCGTCGCGCAGCATGATCTGGC 3720
Qy 1741 AAAAGGTTACTACTCTCAACCCGTGACCCG 1770
    |||
Db 3721 AAAAGGTTACTACTCTCAACCCGTGACCCG 3750

```

```

RESULT 7
AF313916 9359 bp mRNA linear VRL 01-JAN-2002
LOCUS Hepatitis C virus polyprotein mRNA, partial cds.
DEFINITION AF313916
ACCESSION AF313916
VERSION AF313916.1 GI:18027684
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
           Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
           Hepacivirus.
REFERENCE 1 (bases 1 to 9359)
AUTHORS Fanning, L.J., Itakura, J., Nagayama, K. and Enomoto, N.
TITLE Characteristics of Hepatitis C viral genome associated with disease
         progression in a homogeneous patient population
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 9359)
AUTHORS Fanning, L.J., Itakura, J., Nagayama, K. and Enomoto, N.
TITLE Direct Submission
JOURNAL Submitted (17-OCT-2000) Medicine, National University of Ireland,
         Cork, Hepatitis C Unit, Clinical Sciences Building, Cork University
         Hospital, Cork, Ireland
         Location/Qualifiers
FEATURES
source 1..9359
        /organism="Hepatitis C virus"
        /mol_type="mRNA"
        /db_xref="taxon:11103"

```

5'UTR
CDS

/note="Isolated from source plasma identified as the single source of the iatrogenic infection of Rhesus negative women in Ireland from May 1977-November 1978; exposure was through HCV 1b contaminated anti-D immunoglobulin"

1..329
330..9359
/product="polyprotein"
/protein_id="AAL55821.1"
/db_xref="GI:18027685"

/translation="WSTINPKFQRTKANTNRPDQVKEPQGQIVGVYLLIPRGPRLGVARTKTSERSQPRGRQPIPKARPEGRTMAQPGVWPVNGMGMAWMLSPRGSRPQGPDPERRRNLKVIDTLCTGADLMGIPIVGAFLGAALAGVLEADGVNATGNTLPGCSRSFILALISCTIPASILEVRNYSVGVHVNDCSNALIVYEAADMIHMPGCVPCVRENNSRCWVALPTPLTANSSILPTTIRRHDLIVGTAPCSAMYVGDLCGVFLVQLTPGSRHEVVDNCISITPGHNSGHRMADMMNMSPTTALVYVQLIRIPQAVVDVAGAHVGLAGLAYSMAGNNAKVLIVMLFAGVDEGTYTGAQAHTRGLATLFSFSPSONIQLINTGSMHINRLNCNDSLNTGFLAIPARHNASGCPERMASCRPIDRFAQMGPIAYAKPISLDQKPCYMHYAPQCGIVASQVCPVYCFTPSPVVGTTRDFRGVPTYSWNGENETDVLNTRPRPGNMFGCTMMNGTPTCCGPPCDIGVGNDTLICPTDCPKHPEATYTKCGSPWLTPRCMDVYPRILMHPCTVNFITIKRYTVGVEHRLNAAKMTGRCRDLDEDRSELSPILSTTEKQVLPCEFTTLIPALSTGLIHLRNIVDQYLYGISAAVSAIKKEVYLLEFLFADAVCAKLMMLLIVQEALEMLLVINAAVAGTNGILSFVEFCAMYLKRLVGAAYAFYGVPLLLLLALPPRAYAMDREMAASCAGAVFICGLATLTPHYKFLILIMLOYETITPAEHLQWVPIPLNVRGSRDALILMCAIHPELIFATIKILAILAGLPLMLOAGITRVFVRAHGLIRACVIVREAVAGHYVOMALMRLLALNGVYVHTPLRMAHRCGLADLAIVTPVPSDMETKILITWGADTAACGDIILGLPYSARGREILUAPADLSBQGRALPITAYAOQTRGLGLCIVTSLTRDKXNVEBVDVSTATOSPLATCVNGCVYVHAGTKTLAQKQPIQMTYNTNVDQDVLGMLPARGARLTPCTCGSBDLYLVTHADVTPYRRGDSRSLSPRPVSYLKSSGGPLCPSPAGANGIIRAVCTGVAKAVDPVVESEMETTMRSPFTDNSSPPAVPOTFOYAHILAPITGSGSKSPVAAVAAQGVKVLIVNSVAULTGFGVYMKAHVDNIRITGVRTITGAPITYSTGKFLADGCGGAGVDIICDCGHS TDSTTIGIGTVDQARTAGRLVYLATPDPGAVTPHNTIEVALSNIEITPYGKALPLETKGSHLIFCHSRKCCDELAALKLSGLINAAVAYKRGDVSVIPSGDVVVAIDALMTGTGDFDSVIDCNCVTQTVFSLDPTPTITTTVPDQAVSRQSGRGTRVGRTHGAYRRTPEERPSGMFDSVLCCEYDAGACAWELPAETSVSLRAYLTPRVPVCO DHEFMGVFTGLTHIDAHPLSOTKQNGDPVYVLAOVAITVCARAOAPPEPMDQMKC LIRKPTLHGPTLLVRLGAVONENVLIITPTIKTMMQSAIDLAISTWYVGSGLAALAAVCIITGSGVIVGRIILSGRPAITPDEEVLYOQDENBECKTHLPYIEQOQLAB QFKQKALGLQATKQAEAAAPVESKRALSESTWAKHMMFISGITYLADLSLPLGN PAISLMAFTASITSPLTQHTLLFNILAGVAAQOLAPSAAGAFVAGALGAGVSGI GLGVLDIILAGYAGVAGALVAFKNSGSEPTEDVNLPLAIIISGALVGVCAAILRRHVPGEGAVOMNRLIAFASRGHNVSPTHVVPSPDAARTOLISLPTLITOLK RLIHOMINDECSPTGSSGMLDVMWMTCTVITDPFTWLOSLPLPPLVPSICORGX GVMRGDGIQMTTCCGQOITGHVNGSMRIVGPTCSTNTHGTPPIINAYTTPGCTPSP APNYSRALMRVTAEEVLETVQGDPHYVTGTTDNVCCPCQVPAPEPTEVDGRLRL YAPRCKPLREBEVFOVGLNOYLVSQLPCEPEPDVAVLSMLTDPSHIITAEAKRLI ARGSPSLASSASQLSAPSLKATCTRHSDPADLIEANLIMQEMGNTIRYESN KVVILDSFEPLRAREDEREVSPAEILIRSRKFPRAPIVMPDPYNPLVSMQDPY VPPVHGCCLPPTKAPPIPPRRKRTVILESTVYSSALAEATKTSBESASDSTCT ATAPDPSPDGDAGDSVESNRPPIEGEGEDPDLSDGHSITVSEASSEDVYCCNS YTTWGTALITPCSAAESKLPINALNSILRHNHNVATTSRSASQRRKVTFDRLOVLD LSHVVDLKEEMKAKASTYAKLLPVEBCKLPKPPSASKGPKADRNLSKXVNHIL DNYRDLVLEMETPIDITIMAKNVEFCVQEGEKSPKARLIVPELDLVRVEKALVD LVSTLPOAVNGSSYGFQYSPGQRYEFLVNMKKKNGMGVADYDRCDSTYTESDIYV ESTYVOCCLAPPEAROIRSLTRELTYIGSLTKGNGKNGCVRRAAGVLTSGNTLTCTYKASACRAAKLQDCTMLVCGDDLVIVICESGTXEDASLRFVTEAMTRYSAPRG DDPPEPVDLEITSSCNVSVAHDASSKRYVYLRDPTPLARAAWTAHTHRTORLH GLSAFSLHSYSPGEINRVAACTRLGVLPKREHRRARSYVAKLLSGGGRANCGKYLGVNVRTRTKLKTLPISASQDLDSWFAVGSQGIYHLSBARRRWMLCILLISGVG GITYLIPNR"

ORIGIN

Query Match 89.3%; Score 1580.6; DB 14; Length 9359;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 1649; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

Qy 5 ATGACTTCCAGACGCTGACTCATTCGAGCGCAACCTCTGTGCGCGCATAGATGGCG 64
 |||
 Db 6985 ATGACTCCCGCAGATGCTGACTCATTCGAGCGCAATCTCTGTGCGCGCAGAGATGGCG 7044

QY 65 GGGACATTAACCCCGTGGAGTCAAGAAACAAGTAGTAATCTTGAATCTTTGACCCCG 124
 Db 7045 GGAACATCACCCCGTGGAGTCAAGAAACAAGTAGTAATCTTTGACCCCG 7104
 QY 125 TCCGACGAGAGAGATGAGCGGGAAGTGTCCGTCCCGGAGATCTCTCGGAAATCCA 184
 Db 7105 TTGAGACGAGAGAGATGAGAGGAAAGTATCCGTTCGGGGAATCTCTCGGAGATCCA 7164
 QY 185 AGAAATTTCCACCAAGGATGATCCGCAATGGGACGCCCGATTAACAACCTTCGCTGCG 244
 Db 7165 GGAATTTCCCGAGAGATGCCATATGGGCAAGCGGATTAACAACCTTCGCTGAG 7224
 QY 245 AGTCTCGAAGAGCCCGGAGTAAAGTCCCTCAATGTGATCTGGGTGCCCTCCACCTTA 304
 Db 7225 AGTCTCGAAGAGCCCGGAGTAAAGTCCCTCAATGTGATCTGGGTGCCCTCCACCTTA 7284
 QY 305 CTAAAGACCCCTCTATACCACTCCACGAGAAAGAGACAGTGTCTTGAAGAAATCCA 364
 Db 7285 CTAAAGACCCCTCTATACCACTCCACGAGAAAGAGACAGTGTCTTGAAGAAATCCA 7344
 QY 365 CCGTGTCTTCTGCTGCGGAGCTTGCACAAAGGCTTTGGTAGCTCCGACCGTCCG 424
 Db 7345 CCGTGTCTTCTGCTGCGGAGCTTGCACAAAGGCTTTGGTAGCTCCGACCGTCCG 7404
 QY 425 CCGTGCACAGCGGACGAGCAAGCGCCCTCTGACCAATCTTCCGACAGCGGAGCAG 484
 Db 7405 CCGTGCACAGCGGACGAGCAAGCGCCCTCTGACCAAGCGCCCTCCGACAGCGGAGCAG 7464
 QY 485 GATCTACGTTAGTGTATCTCTCATGCGCCCTTGAAGGAGGAGCGGAGGAGCCCG 544
 Db 7465 GATCTACGTTAGTGTATCTCTCATGCGCCCTTGAAGGAGGAGCGGAGGAGCCCG 7524
 QY 545 ATCTCAGCGACGGGCTTTGGTCTACCGTGAAGAGAGCGGAGTGAAGAGTGTCTGCT 604
 Db 7525 ATCTCAGCGACGGGCTTTGGTCTACCGTGAAGAGAGCGGAGTGAAGAGTGTCTGCT 7584
 QY 605 GCTGCATGTCTTCAATGACAGAGCGGCTTGTATCAAGCCATGCGCTCGGAGAGAAAGCA 664
 Db 7585 GCTGCATGTCTTCAATGACAGAGCGGCTTGTATCAAGCCATGCGCTCGGAGAGAAAGCA 7644
 QY 665 AGTGCACATCAAGCGGCTTGAAGCACTTTGTGCGTCAACAATGATGTCTAAGCTTA 724
 Db 7645 AGTGCACATCAAGCGGCTTGAAGCACTTTGTGCGTCAACAATGATGTCTAAGCTTA 7704
 QY 725 CCAATCTCCGACGAGCAAGCGGACGAGAAAGGTCACTTTGAAGAGAGTCAAGTCAATCC 784
 Db 7705 CCAATCTCCGACGAGCAAGCGGACGAGAAAGGTCACTTTGAAGAGAGTCAAGTCAATCC 7764
 QY 785 TGGACGATCACTACAGAGAGTGTCTCAAGAGATGAAGGCGGTCACAGTTAAG 844
 Db 7765 TGGACGATCACTACAGAGAGTGTCTCAAGAGATGAAGGCGGTCACAGTTAAG 7824
 QY 845 CTAAAGCTTCAATGATGAAGAGCGGCAAGGTGACGCCGCCACATTTGGGCAAAATCTTA 904
 Db 7825 CTAAAGCTTCAATGATGAAGAGCGGCAAGGTGACGCCGCCACATTTGGGCAAAATCTTA 7884
 QY 905 AATTGGCTATGGGAGAAAGAGCTCCGAAACCTATCCAGCAAGGCCATTTAACACATCC 964
 Db 7885 AATTGGCTATGGGAGAAAGAGCTCCGAAACCTATCCAGCAAGGCCATTTAACACATCC 7944
 QY 965 GCTCGGTGGAGAGACTTGTGGAAGACACTGAACCAATTTGACACCAATCATATG 1024
 Db 7945 TCTCCCGTGGAGAGACTTGTGGAAGACACTGAACCAATTTGACACCAATCATATG 8004
 QY 1025 CAAAAAATGAGTTTCTGCTGCTCAACAGAGAGAGGCGCAAGCCAGCTCGCTTA 1084
 Db 8005 CAAAAAATGAGTTTCTGCTGCTCAACAGAGAGAGGCGCGAAGCCAGCTCGCTTA 8064
 QY 1085 TGTGTTCACAGCTTGGGGGTCTGTGTGCGAGAAATGAGCCCTTATGACGTGCT 1144
 Db 8065 TGTGTTCACAGCTTGGGGGTCTGTGTGCGAGAAATGAGCCCTTATGACGTGCT 8124

QY 1145 CCACCTCCCTGAGGCTGTGATGGGCTCTCGNAGAGATTCAGATTCCTCTGACAGC 1204
 Db 8125 CCACCTTCCTGAGGCTGTGATGGGCTCTCTCATGAGGTTCCAGTACTCTCTGACAGC 8184
 QY 1205 GGGTCAGTTCCTGTGAAAGCGCTGGAATCAAGAGAGACCCCTATGAGGCTTTGATATG 1264
 Db 8185 GGGTCAGTTCCTGTGAAATGCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8244
 QY 1265 ACACCGCTGTTTGTGATCAAGATCACTGAGATGACATCCGTGTAGAGAGATCAATTT 1324
 Db 8245 ACACCGCTGTTTGTGATCAAGATCACTGAGATGACATCCGTGTAGAGAGATCAATTT 8304
 QY 1325 ATCAATGTTGATGCTTGAGCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1384
 Db 8305 ATCAATGTTGATGCTTGAGCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8364
 QY 1385 TTTATATCGGGGCTCCCTGACTAATTCAGAAAGGAGAGAGAGAGAGAGAGAGAGAGAG 1444
 Db 8365 TTTATATCGGGGCTCCCTGACTAATTCAGAAAGGAGAGAGAGAGAGAGAGAGAGAGAG 8424
 QY 1445 GGGCGAGCGGCTGTGAGAGACTAGCTGCTGATATCCCTCAATGTTACTTGAAGGCTT 1504
 Db 8425 GGGCGAGCGGCTGTGAGAGACTAGCTGCTGATATCCCTCAATGTTACTTGAAGGCTT 8484
 QY 1505 CTGACGCTGTGAGAGCTGAG 1564
 Db 8485 CTGACGCTGTGAGAGCTGAG 8544
 QY 1565 TTGTGTTATCTGTGAG 1624
 Db 8545 TTGTGTTATCTGTGAG 8604
 QY 1625 CGAGGCTATGACTAGTACTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1684
 Db 8605 CGAGGCTATGACTAGTACTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8664
 QY 1685 TGGAGTTGATATCATATGATCTCTCAATGATGCTGCGGAGAGATGATCGGCAAA 1744
 Db 8665 TGGAGTTGATATCATATGATCTCTCAATGATGCTGCGGAGAGATGATCGGCAAA 8724
 QY 1745 GGGTATACTACTCAAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1769
 Db 8725 GGGTATACTACTCAAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8749

RESULT 8
 AF208024 9373 bp RNA linear VRL 30-MAR-2000
 LOCUS Hepatitis C virus strain MD34, complete genome.
 DEFINITION AF208024
 ACCESSION AF208024
 VERSION AF208024.1 GI:7341102
 KEYWORDS
 SOURCE Hepatitis C virus
 ORGANISM Hepatitis C virus
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
 1 (bases 1 to 9373)
 Nagayama, K., Kurosaki, M., Enomoto, N., Miyasaka, Y., Martuno, F. and Sato, C.
 TITLE Characteristics of hepatitis C viral genome associated with disease progression
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 9373)
 AUTHORS Nagayama, K., Kurosaki, M., Enomoto, N., Izumi, N. and Sato, C.
 DIRECT SUBMISSION
 JOURNAL Submitted (24-NOV-1999) Second Department of Internal Medicine, Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-ku, Tokyo 113-8519, Japan
 FEATURES
 source
 1..9373
 /organism="Hepatitis C virus"
 /mol_type="genomic RNA"
 /strain="MD34"

gene
CDS

```
/db_xref="taxon:11103"  
/note="from a patient with persistently normal ALT values  
(ASC-14)"  
1..9373  
/gene="MD34"  
330..9356  
/gene="MD34"  
/codon_start=1  
/product="polyprotein"  
/protein_id="AA061205.1"  
/db_xref="GI:7341103"
```

```
/translation="MSINPKPQKTKRNTNRPDVKKPGSQIVGVYLIDRPRRL  
GVRATKTSERQPGRRQPIKARQPEBRARAOPEYMPILVGNMGWAGMLSPRG  
SNVNGPTDPRRRSRNLKVIDLTCTGFDLMNGYIPVGAAPGABARALAHVRVLD  
GNYATNGPGCSFSIFLLALSLCTIPSAVEVRNVSQVYTVNDSSISIVYAAD  
MIMSPGVCVPCREGNRSQWALPTLAAKRGSVPTTIRRHVDLVGAAGAAFCAMV  
VDLGGSVFVLSQLPFSPRRHETVODCNCSIYPGHSRHRANMNMNSTALLV  
SOLLRIPOAVMDVMAHMGVLAGLALYSIVGNMAKVLIVMLFAGVDMILGSSQD  
TTRRTSLSPSPSQNIQLVKTGMNHNIRALNCNDSHQTEFLAIPYAKHINTSGC  
LERLTSCHPIDKFDQGMGPLYVEHSGSQKQKCMHWYAPPCGIPASVVCVPCFT  
PSPVVGTRDFRGAFTYWGNETVLLNTRPQGMFCGTMMNSSGYTKCGGP  
CNIGVGNDTLICPTDCPKHPEATYTKGSGPMLTPRCMDVYPIRMHYPCVNPV  
PKVRYGVGVEHRLSAACMTRGRCNLEDRSELSPILSTTMOILPCSFLLPA  
ISTGLIHQNIIVDQYLYGSAVVSFTKMEYVLLFLDLADARICACLMWLLIP  
QAEADLENLVLEMAASVAGAHGILSPVFFCAWYIKGLVGAAYALYGPPLLL  
IALPPRAYMDREMAASCGAVFVGLVLTLSPPYKVFARLIMNSOYEITAEALQ  
VWVPLVNGGRDAIILLCAVHELIPTITKILAILGLPLMWLASIIKVYFPAQ  
GLIRACMLARKVAGHYQMAFMKLAALTGVVDHLPLDQMAASIRLADVAEPV  
VSDMETKIITWGADPTACGDIILGLPYVARGKEIILGLDQMSQMRILAPITAY  
SQOTRGLACITTSITGRDRKNGEVOVSTATOSPLATCNGCCTVFRHAGSKTL  
AGPKPIYDMYTNVQDLVGMQAPPGARSLTCTGSSDLYLVTHADVIPRRGDS  
RGLSLPRVSYLKGSSGGLPCSGHAGVIFRAVCTRGKAVADVPEVSEMTLGR  
SEVFNDNSPPAVPOTFOVAHILAPTSKSTKVAAYAAQYKLVINPVSVALTGF  
GAYMKKAGVDPNIRTGARTVTGAFTYTSYKFLADGGCGGAYVIDIMCECHSTD  
STSIIGTIVLDOAETAGARLVLAATPRGSVTVPHPIEVALSNTEGPEPKKAI  
PIETIKGRHLIFCHSKKCDLAKLSGLGNNAVAYTRGDEVSTPASGNVVAID  
AMTGTGDFSDVIDCNTCVITVPSLDPTFTTETTVPDVASRSQRKGTGRR  
GIYREVTPEGRSPGMDSSVLCECYDAGCAWELPAETVSLRAYLPLTGLPVCODH  
LEFWSVFTGLTHIDAHPLSQKQADNPYLAOATVCAARAPSPMDMCLIL  
RLKPLTGGTPLLRLGAVONENVLTHTPTKXIMACMSDLESHVTVLGVGVAL  
LAAYCLITGSSVIVGRILISGRPAVIPDRVTVREFPDECHSPYIEQMOLEOP  
KOKAGLIOTATKQAEVAPVYESKRALBAFMAKMMNFISGITYLAGSLTPENPA  
IASLAFSTASITSPITOSTLFLNLLGGVAAQLAPPSLASVAGIAGAASVIGL  
GVLVIDIAGYAGVAGALVAFKNGSEMPSTEDLVNLPALISFAGIAGVAGVAGL  
RHVPGEBGAVQMMARLIAFASGNHVSPTHVPESDAAARVQISSLITTIQKRL  
HOWINDECSPTCSGSLMDPMDMICTVLDFTKWLQSLAPLPPLPSPFSOCRGYK  
WEGDGIQMTTCGCAOITGHNVSMSRLIGPKTCNTWHTGTPINAYTTGPTSPAP  
NYSRALMRYAEIEYVTRVGDVPHYVTGTTDVAKPCVPAPEFTFVDGRHLRYA  
PCKPLREEVTFQVLANQYLVGSQLPCPEPBDVAVLTSM.LDPSHITAEAKRRLAR  
GSPPLASSASQLSAPLSKATCTTHHSPDADILEANLPMOENKGNITRVESNKY  
VLDSFDPLRAEBDEREVSPALILKRSKFPAMPIMARPYNPPLLESMDPYVP  
PVHGCPLPPTKAPPIPPRRKRTVLTSTVSSALABLATPTSSSSAAVDSGTAT  
APDQSPDSDAGSVESYSSMPLLEGSGDDPLSDGMSYTSSEASDVCNSYAT  
WTGALITPCAAEGSKLPIYASNSILRHNNYATTSASNRQKQKVFDRQLQVLDH  
YRDLKEMKAKASTVAKULSVEACKLTPHSASRKGYGADVRNLSCEKALNIRS  
VKDLLEBTEPITDITTIKAKNBIFCVQEPKGRKPARLIVFDLDRNLSCEKALNIRS  
STLPQWASSYGFQYSPQRYEPLVDMAMSKKCMGAVRCDVSTYEDNIVIEE  
SYVCCDLAPSEARQAIRSLTERLYVGLPINSKQNGCRRRACGAVLTSQGNLTLC  
YIKASACRAKLRDCTMLVCGDDLVVICESAGQEDPASAPPEFMEARTVYASGDP  
PYKREYDLBITSCSSNVSVAHDSKRYVLTFRDPTPLARAAWETASTPYNSMNGN  
IIMVAPTLMARILMTHPFSLIAQESKALDCQIYACVSIIEPLDLPQIILQRLHGL  
SAFSLHSPGSEINRVASCLRLGVPLVWRHARSYAKLLSGGGAALCGKTLFN  
WAVRTKLKLTPIPAASQDLSNMFWAGSGDIIYHLSRARPRMFLCLLLSVGVI  
YLPNR"
```

ORIGIN

Query Match 88.7%; Score 1570.6; DB 14; Length 9373;
Best Local Similarity 93.0%; Pred. No. 0;
Matches 1645; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 1 CAANAATGACTTCCCGAGCGCTGACCTCATCGAGGCCAACCTCTGTGGCGGATGATGATG 60
DB 6975 CATCATGACTCCCGAGCGCTGACCTCATCGAGGCCAACCTCTGTGGCGGAGAGATG 7034

```
QY 61 GCGCGGACATTTACCGCGTGAGTCAGAGAACAAAGTAGTAATCTGGACTCTTCGAC 120  
DB 7035 GCGGAAACATCACCCCGGTGAGTCAGAGAAATAGTAATCTGGACTCTTCGAC 7094  
QY 121 CCGCTCCGAGCGGAGAGATGAGCGGGAAGTGTCCGTCCCGCGGAGATCTTCGGA 180  
DB 7095 CCGCTCCGAGCGGAGAGATGAGCGGGAAGTGTCCGTCCCGCGGAGATCTTCGGA 7154  
QY 181 TCCAGAAATTTCCACCGAGCATGCGCGCATGCGGCGGANTTAAACCTCCGCTG 240  
DB 7155 TCCAGAAATTTCCCGCGCATGCGCATGCGGCGGANTTAAACCTCCGCTG 7214  
QY 241 CTGAGCTCTGAAAGGCGCCGGAATCACTCCCTCAGAGTATAGTGTGCGCATGCA 300  
DB 7215 CTGAGCTCTGAAAGGCGCCGGAATCACTCCCTCAGAGTATAGTGTGCGCATGCA 7274  
QY 301 CCTACTAAGACCCCTCTATATACACCTTCACGAGAAAGAGACAGTTGTTCTGAC 360  
DB 7275 CCTACTAAGACCCCTCTATATACACCTTCACGAGAAAGAGACAGTTGTTCTGAC 7334  
QY 361 TCCACCGGTCTTCTGCGCGGAGGTCGCAAAAGCTTTGTAGTCTCGGAGCG 420  
DB 7335 TCCACCGGTCTTCTGCGCGGAGGTCGCAAAAGCTTTGTAGTCTCGGAGCG 7394  
QY 421 TCGGCGTCCGACAGCGGACCGGCAACCGCCCTCCTGACCAATCCTCCGACG 480  
DB 7395 TCGGCGTCCGACAGCGGACCGGCAACCGCCCTCCTGACCAATCCTCCGACG 7454  
QY 481 CGAGATTCGACGTTGAGTGTATATCTTCATGCCCCCTTGAAGGGAGCGGGAG 540  
DB 7455 CGAGATTCGACGTTGAGTGTATATCTTCATGCCCCCTTGAAGGGAGCGGGAG 7514  
QY 541 CCCGATCTCAGCGAGCGGCTCTTGCTACCGTGAAGTGAAGGCGCGAGAGAG 600  
DB 7515 CCCGATCTCAGCGAGCGGCTCTTGCTACCGTGAAGTGAAGGCGCGAGAGAG 7574  
QY 601 TGTGCTGATGATGCTCTACATGAGCAAGCGCTCTGATCAAGCATGCGTCCGAG 660  
DB 7575 TGTGCTGATGATGCTCTACATGAGCAAGCGCTCTGATCAAGCATGCGTCCGAG 7634  
QY 661 AGCAAGCTGCCCATCAACGCGTTGAGCACTCTTGTGCTCACACACATGCTTAC 720  
DB 7635 AGCAAGCTGCCCATCAACGCGTTGAGCACTCTTGTGCTCACACATGCTTAC 7694  
QY 721 GCTACCAATCCCGAGCGCAAGCGAGCGGAGAAAGGTCACTTTGACAGACTGAA 780  
DB 7695 GCTACCAATCCCGAGCGCAAGCGAGCGGAGAAAGGTCACTTTGACAGACTGAA 7754  
QY 781 ATCTGACGATCACTACAGAGAGTGTCTCAAGAGATGAAGGCGGAGCTCCACAG 840  
DB 7755 ATCTGACGATCACTACAGAGAGTGTCTCAAGAGATGAAGGCGGAGCTCCACAG 7814  
QY 7815 AAGGCTAAGCTTTATATCATGATGAGAGAGCTGCAAGCTGAGCGCCCACTTC 900  
DB 841 AAGGCTAAGCTTTATATCATGATGAGAGAGCTGCAAGCTGAGCGCCCACTTC 900  
QY 901 TCTAATTTGGCTATGAGGCGCAAGAGAGTCCGGAACCTATCAGCAAGGCCATTAA 960  
DB 8745 AAGGCTAAGCTTTATATCATGATGAGAGAGCTGCAAGCTGAGCGCCCACTTC 7874  
QY 961 ATCCGCTCCGTGTGGAGAGCTTGTGGAAGACATGCAACCAATTTGACACACATC 1020  
DB 7935 ATCCGCTCCGTGTGGAGAGCTTGTGGAAGACATGCAACCAATTTGACACACATC 7994  
QY 1021 ATGCAAAAAATGAGTTTCTGCGTCCACACAGAGAGAGCGCCGACGACTCGC 1080  
DB 7995 ATGCAAAAAATGAGTTTCTGCGTCCACACAGAGAGAGCGCCGACGACTCGC 8054  
QY 1081 CTTATCGGTCCCAAGCTTGGGGGTCCGTGTGCGGAGAAATGCGCCCTATAGAGTG 1140  
DB 8055 CTTATCGGTCCCAAGCTTGGGGGTCCGTGTGCGGAGAAATGCGCCCTATAGAGTG 8114
```



```

Qy 1141 GTCTCAACCTCCCTCAGGCTGTGATGGGCTCTCGTAGCGATTCCAGTATTCCTCGGA 1200
Db 8115 GTCTCAACCTCTCTCAGGCGGTGATGGGCTCTCGTAGCGATTCCAGTATTCCTCGGA 8174
Qy 1201 CAGGCGGTGAGTTCCTGTGTAAGCGCTGTGAATAAAGAACCCCTATGGCTTTGCA 1260
Db 8175 CAGGCGGTGAGTTCCTGTGTAAGCGCTGTGAATAAAGAACCCCTATGGCTTTGCA 8234
Qy 1261 TATGACACCGGCTTTTGAAGTCAAGTCACTGAGAAATGATCCTGTGAGAGAGTCA 1320
Db 8235 TATGACACCGGCTTTTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 8294
Qy 1321 ATTATCAATGTTGTGACTGGCCCGGAGACGAGACGAGTCAAGTCAAGTCAAGTCA 1380
Db 8295 ATTATCAATGTTGTGACTGGCCCGGAGACGAGACGAGTCAAGTCAAGTCAAGTCA 8354
Qy 1381 CGGCTTTATATCGGGGCTCCCTGACTAATTCAAAAGGAGAACTGGCGCTATCGCGG 1440
Db 8355 CGGCTTTATATCGGGGCTCCCTGACTAATTCAAAAGGAGAACTGGCGCTATCGCGG 8414
Qy 1441 TGCGGGCGGAGCGGCTGCTGAGACGACTAGTGGCGTAAATCCTCAGATTTACTTGAAG 1500
Db 8415 TGCGGGCGGAGCGGCTGCTGAGACGACTAGTGGCGTAAATCCTCAGATTTACTTGAAG 8474
Qy 1501 GCTCTGACGCTGTGAGCTGCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 1560
Db 8475 GCTCTGACGCTGTGAGCTGCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 8534
Qy 1561 GGGCTTGTCTTATCTGTGAGACGCGGGAACCCAGAGAGACGCGGCGACTACAGAGTC 1620
Db 8535 GAGCTCGTGTATCTGTGAGAGCGGCGGAGACCCAGAGAGATCGGCGAGCTTACAGAGTC 8594
Qy 1621 TTACGCGAGCTTATGACTAGTACTCTGCCCCCGGCGGAGCCCGCCCAAGCAATAC 1680
Db 8595 TTACGCGAGCTTATGACTAGTACTCTGCCCCCGGCGGAGCCCGCCCAAGCAATAC 8654
Qy 1681 GACTGGAAGTGAATTAACATCATGCTCTCCATGATGTGCGGTGCGGACGATGATCGGC 1740
Db 8655 GACTGGAAGTGAATTAACATCATGCTCTCCATGATGTGCGGTGCGGACGATGATCGGC 8714
Qy 1741 AAAAGGTAATTAATCACTGACCCGTAACCC 1769
Db 8715 AAAAGGTAATTAATCACTGACCCGTAACCC 8743

RESULT 9
AY587845 Hepatitis C virus strain RFL_2k/1b, N687 polypeptide gene, complete
LOCUS AY587845 9357 bp RNA linear VRL 25-JUN-2004
DEFINITION Hepatitis C virus strain RFL_2k/1b, N687 polypeptide gene, complete
cde.
ACCESSION AY587845
VERSION AY587845.1 GI:46560635
KEYWORDS Hepatitis C virus
SOURCE Hepatitis C virus
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepatitis C virus
REFERENCE 1 (bases 1 to 9357)
AUTHORS Kallinina, O., Norder, H., and Magnus, L.O.
TITLE Full-length open reading frame of a recombinant hepatitis C virus
JOURNAL J. Gen. Virol. 85 (Pt 7), 1853-1857 (2004)
PUBMED 15218169
REFERENCE 2 (bases 1 to 9357)
AUTHORS Kallinina, O., Norder, H., and Magnus, L.O.
TITLE Direct Substitution
JOURNAL Submitted (01-APR-2004) Virology, Swedish Institute for Infectious
Disease Control, 171 82 Solna, Sweden
FEATURES
source
1. 9357
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/strain="RFL_2k/1b, N687"

```

```

/db_xref="taxon:11103".
/note="natural recombinant"
<1..268
269..9313
/note="encodes C, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A,
NS5B"
/codon_start=1
/product="polypeptide"
/protein_id="AY00644.1"
/db_xref="GI:46560636"
/translation="MSTNPKPQKTKRNTNRNRRPDVYKFGGQIVGVYLLPRGRRL
GVATRTSERSQWRGRQPIEDRSSTGRSDYVPMPLVGNELGALMAGVLEDSRG
SRSPMGTPTRRSRMLKIVITLTCGLDMLGVIVPAGVPGVAGLAAHGVLEDS
GINVATGNTLPGCSFSTLALISCISSVPAVNTSTQVTAATDNCSSSTLMWLEG
AVLHNGCTPCELVGVNRCWLPVTPNVAVRERAGLTGKRLHIDILIVSATISALY
IDVCGAAMTAAQAVTISPERIAFVDDCSTYIPGHTGHRNAMDMMWSPSTTNV
AYLMRLPEIVLIDITGAMHGVNGLAFPMQMAKAAVYLLTLTAVESSTITMGQA
SHTMGLVSI FSGARONIQLINTGSMHINTALNDSLOTGFIASLPYVRFNS
GCEPERSSCRPIEGPRIGMSLKYEENVTNDEDMRYCWHYKPCGIVPRTVCGPV
YCPPTSPVYVGTDDKYGAPTYGEMENOVPLINSTRPRGAMFGCTMNGGFTKTC
GAPPCIRIRDNSTEDLCTPCFRKRPATYTKGAGFWLTKLVDYRYLMHYP
TANFTFKIRMFIGVEHRLNACNTRDRCNLEDRDSQSLPLHSTTEMAVILCT
YSDMPALSTGLHLHQNIVDVOYLGLSPALFTYKMWVLLPLLDADAVACML
MILLGQAEALAEKLVILHAASAFNGIICFTIFIAAMHILKGRVPLATYSYGLM
SLLILIALPQOAVLNLNTEGOLGLVILVSVLISPAVYILCRSLMLSYLVR
AAALIOEWVPRQAGRGDCITMAVITLCPGVIPDTIKLALVLEGHILRSVTSVP
YFRKALALRTICAVRHLSGKRVQMMILTLKRGTYIYDLTLPQDAHAGPITLV
VAEPVFSDMETKLIITWGAETACDIIILGIPVSRGRREILGPDLEQGMRL
APITAYSQOTRGLGCTVTSLGRBNOYQGVQVSTATOSFLATCVNGVCTVYHG
AGSKTLGAKRPILOMYTNDODLVGMPPAPGARSITPCTGSSDLYLTRADYIPV
RRRGSRGSLSPRISYKSSGGGHLCPSHANGVILPAAICTGVAAVAVFVPS
METTWRSPYFTNNSPRAPVQTPQVAAHLAPGSGKSTVPVAAQOKVLYLVNSV
AATLGRGAYMSRAVGVDPNRTGVRITTGATITTSYKGVPLADGSGCAVDILCD
ECHSDTSTSLIGITVLDQATAGALVVLATPBGSVTPHSPLEVALNIGIP
PYKAKIPIEVILGGRHLIFCHSKKCDLAAKLSALGNAVAYYGLDVSVIPSGDV
VYVADTALMGTGDEDSVIDCNTCYOTVDSLDPTIETTTMPODAVSORGR
TORGGRGIRFPTPGRPSGMSDVLCECVAGAMWELTPAETSVRLAYLVNPLG
PVCODHLEWESVFTGLTHIDAPLSQTKQABENPYLIAYATYCARQAPRPSMD
MMKCLRLKPTLHGPPLVLRGAVONETKPIIDPRLEITKFMACMSADLEVTSWLVG
GVLAALAACTLTGGSVIVGRVLSGKPIIDPRLEITKFMACMSADLEVTSWLVG
OLAEQFKOVALGILHTATKOAEAAA PVESKVRALETFAKMMNFIISQIYVLAGST
LBNPVALISLAFASITSPITHTOHLVLEPILAGVWAAOLAPASASQOKVYTPG
VSGISGKLVYDILAGVAGVAGVAFVPMGEMPTEDVLLPALISPGALVAVG
VGAALIRHVRGEGVAGVOMNRLIAPASGHNVSPTHYBESDAAARVQIISLITV
QILKRLHQNINDECTPSCGSLWDVDMITCTVLDFTKTLQSKLPQLGPRPFSQ
RGYKGVWRGDGIMQTTVPCGAQITGHVKNKSRIVGPKTCSNTWGTPEFINAYTGPC
SPSPALNYSRRLMRVAABEYVEVTVGVGTMTTDDNKPCQVPAPEFTEVDVG
RLLHRYAPACKRLREAVTFOGLNOYLVGSSLPCEBEPVAVLTSMLTDPISILATA
KRLLRGSPPLASSASQLSAPSKATCTTHDDPDDLLIENLILRDEMGNILTRV
ESBNKYVULDSFPIRPREDEREVSYPALILKSKKTPRAMITNAPDPNPLASMK
DDYVAPVHVHGCPLPPTKAPLIPPRKRTVVLTSVSSALAEATKTFGSGSSAV
DSGTASAPONQTSNDGDSVDSYSPMEPEBDDPDLSDGSSVSEBASSEDVVC
CSMSYWTGALTTPCAAEESKLPINALSNLSLHNHNVATTSASASQOKVYTPG
OVLDDHVRVILKEMKAKASTYAKLISVEAKCLTPHSPARKSGFGAGQDVNLSKA
VNHISVWKDILEDIETPTITMAKINVEFCOPKRGKPARLIVPLIDRSPYCEM
ALHYVSTIPQAVMPSYGFQYSPQKRVFELVYNAKSKSKSPNGRAYTRCSTYVEN
DIRVESITQCCDLAPEARQAKSLTERIYIGGFLNSGQNGCRKRCASGVLTTC
GNTLTCYLAASACRAKRLDCMTMLVOCGDVILCESACTODASLRATFETAMRYS
APGPPPOPEYDELELITSCSSNVSAHAGSRRVYLLTDPPTPLARAMEFTRAPYS
NMLGNILNITLWARMILMTHPESILIAQOLKALDCCOYIGCASIEPDLPOI
OHLGLISANSHSVASRGEINRYASCLRIKGVPLPAMHRAVSAYAKLISQGRYATC
GKYLNRNAVTKTLKLTIPPAASQDLDSGHFVAGVSGDITVHLSRARPRMFWCLLL
SVGVGIVLDNR"
9314..>9357

```

ORIGIN

```

Qy 1 CAAATGACTTCCAGACGCTGACTGATCGAGGCAACCTCTGTGCGGATGAGANG 60
Db 6932 CATCATGACTCCCAAGAGCTGACCTCATAGAGGCAACCTCTGTGCGGATGAGANG 6991

```

QY	61	GGCGGGGACATTACCCCGGTGAGTCACAGAACAAAGTAATCTTGGACTCTTTGGAC	120
Db	6992	GGCGGAAACATCACCCGCGTGAAGTCAGAGAAATGAAGTAAGTATTGGACTCTTTTGGAC	7051
QY	121	CCGCTCCAGCGGAGAGAGATGAGCGGGAAGTGTCCGTCCGCGCGAGATCCTCGGAAA	180
Db	7052	CCGATCCGGCCGGAGAGAGACGAGAGGGAAATATCCGTTCCGCGCGAGATCCTCGGAAA	7111
QY	181	TTCAAGAAATTTCCACACGACGATGCTCCGCAATGGGACGCTCCGATTTACAACCTTCGCTG	240
Db	7112	TTCAAGAAATTTCCCGGACGATGCTTATGGGACGCGCGGATTTACAACCTTCGCTG	7171
QY	241	CTGAGTCTTGGAAAGCCCGGACCTACGATCCCTCCAGTGTATCAATGGGTGGCCACTGGCA	300
Db	7172	TTAGGCTCTTGGAGAGACCCGGACTAGTTCATCCGTGTGTGCACGGGTGGCCACTGGCA	7231
QY	301	CCTACTAAGACCCCTCTTATACCACTTCCACGAGAAAAGAGACAGTTGTCTGACAGAA	360
Db	7232	CCTACTAAGACCCCTCTCAATACCACTTCCACGAGAAAAGAGACAGTTGTCTGACAGAG	7291
QY	361	TTCAACCGTGTCTTCTGCTCTGGCGGAGCTTGGCACAAAGGCTTTTGTAGTCCGAGACG	420
Db	7292	TTCAACCGTGTCTTCTGCTCTGGCGGAGCTTCTAAGAGACCTTTTGTAGTCCGAGATCG	7351
QY	421	TCGGCCGTGACAGCGGCGACGAGCAACCGCCCTCTGACCAATCCTCCGACGACGCGAGA	480
Db	7352	TCGGCCGTGACAGCGGCGACGCGCGCTCCCTCAAAACCAAGACTCCAAAGCGGCGAC	7411
QY	481	GCAGGATTTGACGTTGAGTCTGTAATCTTCAATGCCCCCTTTGAGGGGAGACCGGAGGAC	540
Db	7412	ACAGGATTCGAGCTTGAAGTCTGTAATCTTCAATGCCCCCTTTGAGGGGAGACCGGAGGAC	7471
QY	541	CCCGATCTCAGGAGACGGGTCTTGGCTTACCGTGAAGTGAAGGCGCGTGAAGAGATGCTC	600
Db	7472	CCCGATCTCAGGAGACGGGTCTTGGCTTACCGTGAAGTGAAGGCGCTAGTGAAGAGATGCTC	7531
QY	601	TGCTGCTGATGTCTTACACATGACAGAGGCTCTGATCACGCCATGCGTGCAGAGGAA	660
Db	7532	TGCTGCTGATGTCTTACACATGACAGAGGCGCTCTGATCACGCCGTGCGCGGAGGAA	7591
QY	661	AGCAAGTGGCCCATCAAGCGGTGAGCAACTTTTGTGCGCTCACACCAATAGTCTAC	720
Db	7592	AGCAAGTGGCCCATCAATAGCGTGAAGCAACTTTTGTGCGCTCACCAATATGCTCTAT	7651
QY	721	GCTACACATCCCGAGCGCAAGCGACGCGGAGAGAAAGTCACTTTGACAGACTGCAA	780
Db	7652	GCCAACAACATCCCGAGCGCACCAAGCGGAGAAAAGTCACTTTGACAGACTGCA	7711
QY	781	ATCTTGAACGATCACTTACAGAGCGTGCTCAAGAGATGAAGGCGAGGCGTCAAGTT	840
Db	7712	GTCCTGAGACGATCACTTACCGGACCGTGTCAAGAGATGAAGGCAAGGCGTCAAGTTC	7771
QY	841	AAGGTAAGCTTCTATCAGTAGAGAGACCTTGAAGCTGACGCCCCCAATTCCGGCCAAA	900
Db	7772	AAGGCAACCTTCTATCCGTGAGAGACCTTGAAGCTGACGCCCCCAATTCCGGCAGA	7831
QY	901	TCTAATTTGGCTATGGGCAAAAGGACGTCGGAACCTATCAGCAAGGCCATTAAACAC	960
Db	7832	TCTAATTTGGCTATGGGCGAAGGACGTCGGAACCTATCAGCAAGGCCATTAAACAC	7891
QY	961	ATCCGCTCCGTGTGGAGAGACTTGTGGAAAGCACTGAACAACCAATTGACACCAACATC	1020
Db	7892	ATCCACTCCGTGTGAGAGACTTGTGTGAAGCAATTGAACAACCAATTGACACCAACATC	7951
QY	1021	ATGGCAAAAATGAGTTTCTGCTCCCAACAGAGAGAGAGGCGCAACCTCACTGCG	1080
Db	7952	ATGGCAAAAATGAGTTTCTGCGTCCAGCCAGAGAGAGGCGCAACCTCACTGCG	8011
QY	1081	CTTATCGTGTCCAGACTTGGGGGTCCGTGTGTGCGAAGAAATGCGCTCTATGACGTG	1140
Db	8012	CTTATCGTATCCAGATTGGGGGTTGTGTGTGCGAAGAAATGCGCTCTTATGACGTG	8071

OY		1141	GTCCTCACCCCTCCTCGACGGCTGTGATGGAGCTTCTTGTAACGAATTCCAGATATTCTCTGGA	1200
Dd		8072	GTCTCCACCCTCCCTCGACGGCGGTATGGGCCCCCCTCATACGGAATTCCTCTTGGG	8131
OY		1201	CAGCGGGGTGCAGTTCTCTGTGTAAAGCGCTCGAATAACAAGAAGACCCCTATGGGCTTTGCA	1266
Dd		8132	CAGCGGGGTGCAGTTCTCTGTGTAAAGCGCTCGAATAACAAGAAGACCCCTATGGGCTTTGCA	8191
OY		1261	TATGACACCCGCTGTTTTGACTCAACAGTCACCTGAGAAATGACATCCGTGTAGAGAGTCA	1320
Dd		8192	TATGACACCCGCTGTTTTGACTCAACAGTCACCTGAGAAATGATATCCGTGTAGAGAGTCA	8251
OY		1311	ATTATATCAATGTGTGACTTTGGCCCCCGGAAGCCGACAGCGCATTAAGTCCCTCACAGAG	1380
Dd		8252	ATTATATCAATGTGTGACTTTGGCCCCCGGAAGCCGACAGCGCATTAAGTCCCTCACAGAG	8311
OY		1381	CGGCTTTATATCGGGGGGTCCCTGTAAGTAATCAAAGGGGAGAACTGGGTATCGCCGG	1440
Dd		8312	CGGCTTTATATCGGGGGGTCCCTGTAAGTAATCAAAGGGGAGAACTGGGTATCGCCGG	8371
OY		1441	TGCCGCGGAGACGGCGCTGTGACACGACTAGTCGGGTATACCTCAATGTTACTTTGAAG	1500
Dd		8372	TGCCGCGGAGACGGCGGTGTGTGACGACTAGTCGGGGACAAACCTCAATGTTACTTTGAAG	8431
OY		1501	GCTCTTGACGCTGTGTGACGCTGCAAGTCCAGACTGACGATGCTGTGTGCGAGAC	1560
Dd		8432	GCTCTTGACGCTGTGTGACGCTGCAAGTCCGGGAGCTGACGATGCTGTGTGCGAGAC	8491
OY		1561	GGCTTGTGCTTACTGTGTGAGAGGGCGGGAAACCAAGAGAGACGGGGGAGCTTAGCGAGTC	1620
Dd		8492	GACCTTGTGCTTACTGTGTGAGAGGGCGGGAAACCAAGAGAGACGGGGGAGCTTAGCGAGTC	8551
OY		1621	TTCAACGAGGCTATAGACTAGACTAGTCTTGCCCCCGGGGAGACCCGCCCAACAGAAATAC	1680
Dd		8552	TTCAACGAGGCTATAGACTAGACTAGTCTTGCCCCCGGGGAGACCCGCCCAACAGAAATAC	8611
OY		1681	GACCTGAGATTGATPAACATCATCTCTCCAAATGTGTGTGTCGGCACGATGATCTTGGC	1740
Dd		8612	GACCTGAGATTGATPAACATCATCTCTCCAAATGTGTGTGTCGGCACGATGATCTTGGC	8671
OY		1741	AAAAGGGTATCTACCTCAACCCCGAGACC	1769
Dd		8672	AAAAGGGTATCTACCTCAACCCCGAGACC	8700
RESULT 10				
HPCXR2				
LOCUS			Hepatitis C virus (strain HCV-1b, clone HCV-KI-R2), complete genome sequence.	
DEFINITION				
ACCESSION			D50481.1 GI:1030705	
VERSION			D50481	
KEYWORDS			polyprotein; core protein; E1; E2; NS2; NS3; NS4A; NS4B; NS5A; NS5B; envelope protein; non-structural protein; interferon-sensitive; interferon-resistant; IFN-sensitive; ISDR; interferon sensitivity determining region; HVR; hypervariable region.	
SOURCE			Hepatitis C virus	
ORGANISM			Hepatitis C virus	
			Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.	
REFERENCE				
AUTHORS			Enomoto,N., Sakuma,I., Asahina,Y., Kurosaki,M., Murakami,T., Yamamoto,C., Izumi,N., Marumo,F. and Sato,C.	
TITLE			Comparison of full-length sequences of Interferon-sensitive and resistant hepatitis C virus 1b. Sensitivity to interferon is conferred by amino acid substitutions in the NS5A region	
JOURNAL			J. Clin. Invest. 96 (1), 224-230 (1995)	
MEDLINE			95340824	
PUBMED			7542279	
REFERENCE			2 (bases 1 to 9410)	
AUTHORS			Enomoto,N.	
JOURNAL			Unpublished	

REFERENCE
AUTHORS
TITLE
JOURNAL

3 (bases 1 to 9410)
Enomoto, N.
Direct Submission
Submitted (08-MAY-1995) Nobuyuki Enomoto, Tokyo Medical and Dental
University, Second Department of Internal Medicine, 1-5-45 Yushima,
Bunkyo-ku, Tokyo 113, Japan (E-mail: PYN04522@niftyserve.or.jp,
Tel: 03-3813-6111 (ex.3224), Fax: 03-3818-7177)

FEATURES
source

Location/Qualifiers
1. 9410
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/strain="HCV-1b"
/db_xref="taxon:11103"
/clone="HCV-K1-R2"
/note="IFN-resistant HCV from patient 2"
330..9362
/codon_start=1
/product="polyprotein"
/protein_id="BAA09072.1"
/db_xref="GI:1814086"

CDS

/translation="MSTNPKPQKTKRANTNRPRQDYKFGCGGQIVGCVYLIPRGRRL
GVRATKRTBSRQGRKQPIPAQRPBGMAAQPEYPMPLVGENEGMAGWLLSRG
SRSPGPTDPRRSNMGKVIDLTGCGADLMGIVLGAFLGASRALAHGRVLRSD
GNVATGNLPGCSFSIFLLASCLTI PASAEVRNAGVHYVNTDSCNASIYEAD
MIMHTGCVPEANSSRCWVALPTLAAENSVPTTIRHVLLVGAARCSAMY
VVDLGGSVLVQLPTSPRRHETDQDCSIRGHSGRHAMMMMMSPPTALV
SOLIRIQAQVMDVAGAHGVTLAGLAYSVMGMAKYLVMLEPAVDGNTYVSGAK
SHTQGLVSLFPAQAQKIQLVNTGSMHINRTALCNDLSLNTGLFALPTYHRESS
GCEERLASCRSIDAFQGWGPIYVABGSDORPYCMHYAPRCQIVLAPYHRESS
FTSPVVGITDSDGAPTYSGENETVLLNTRAPQSGMFCSTMMNGTGFETGCG
PCNIGVGNTLTCTDCEPRKHEATYKCGSPMLTPTRCWDYEMLYHNYCTFNP
TIFKVMVYGVEHRLAANMTRGERCLEDRSLSPLISTEMOILPESFTPL
PIVSTLHILHONIVDQYLYGSAVSVFYKWTLLFLILADARCAQCLMMML
LQAQEALENLVNLAASVAGAGFLSPVFCAMVYIKGLVPAVAIVYGMVPL
LIALPRAVYAMDREMAASCGAVFGVALLSPHYKVLVALIMVLYFTIRACAP
AQLWIPLVNRCGRDAIILTCALHPELIFSTIKILVALIGPLMVLAQGITPKPYFR
ACGLIRACMLVRGAGHYVOMAFMLALITGYVDHLPTLDMATGRLDAVVE
PVESDMETKIITWGDPTAACDILIGLASARKEITPCTGSEGGGGLAAIT
AYSQGTREGLGCTITSLTGRDKQVEGEQVYSTANOSFELATCTGSCVTYHGAISK
TLAAGKPIITQMTNVNDQLVGMQAPSASLPTCGSSDLTVLRHADVLPVRRG
DGRGSLSPRPVSYLKGSSGGLCPGSAHVGIFRAVCTGAGKAVDIPIESMATT
MSPVFTDSSPPAVPQTFQVAHAPTSKSGSKNPKVEVAGAYKVLVNPVATL
GGAWSKAGWDPNIRGTGRTITGGAATYTSYGVKFLADGGGSGGAGIILICDEGHS
TGSTSLIGITVDQAEETAGARLVYATATPREGSVYVPHNIEVALNSTGERIPGK
ALPITIKGRHLIFCHSKKCDLAKLSGGINAVYRGIDVSI PASGVVVA
TQALMTGFTGDPDVIDCNVCTVQVDFSLDPTLETITVPDASRQQRRTGRG
RAGIYRFTVTPGERSPGMPDSVLCCEYDCAMAYELTPTSTRLRAYINTPGLPVWQ
DLHEFVESFTGLTHIDAPHSOTKOAGNFPYLVAYQACARAAQPPSMQMKC
LIRKFTLHGPPLLYRLGAVONEVILTHPTIKYIMACMSADILEVTSMTLVGGLA
ALAYCLTIGSVIVYGRITILSKPAILPREVLYREFDEMEBASHLPTIBQMOLAE
QPRQKGLLQRTIKQAEAAAPVESKQQLLETFMKHMNFI SGIQYLAGLSTLFGN
PAIASLMAFTASTISPLTQHTLLFNILGVAQAOLA PPAASA PQRKATPDLQVLD
GAGKVLVDILAGAGVAGALVAFKVMGSPSTEDLVLPILPGLVAVGVCAA
ILRBVGPBGAVOMNRLIAFASRGNHSPTHVPESDAAAVTIIISLITLTK
RLHOWNEDCSTPSCGSMLRDWDMTCTYLSDEKTMVLSRVLPRLEGVPLSGORVYK
GWRMGDINGMOTTCPCGAQITGHAKNSMAYVGRKTSNTHRGVPLINATITGCTSP
APNYSALMRVAABEYEVTRGVDFHYVGMTHDKPCQVYPAFFEFELDEVRILR
YAPASKPLLRDEVTQVGNQVYVGSQCEPEPDVAIVTSMLTDSHLYAEAKRL
AAGSPSLASSSASQASPLKATCTTHDSPADLIEANILMRQMGNTITVESEN
KVILDSFEPRLAEDEDEREVLPAELIRSKRPRAMPILARDVNPPLIESGNDY
VPPVHGCPLPTKAPRIPRPRKRTVYLTESTVSALAEIKTKTGSSESSAASDGT
ATAPRPPSSDDAGSDVESMPLEBEGDPDLSDSKMSVSEBAEDVYCCMS
YMTGLITPCAEBSKPLINALSNSLRLHNNHNAVITRSVSQRKATPDLQVLD
DHRYDLKEMKAKASTVKARLTSVEACKLTTPHARSFKYGVAQNRVLSLAINH
RVMXDLQEDTEPTDITIMAKNEVFCVQPEKGRPARLIVPDLGVCEEMALYD
VSTLPOAVGASVGYOYSGRQREVEFLVNASKKCPMGAYRATCFEDSVENDLRV
ESISVOCCLAPBAROIRSLTERLYIGPLTNSKONGYRARSCTSGVLTNSCGNTL
TEYLKRTAACRAKQDCTMLVCGDVLVYCSAGTQEDBASIRVSTAMTRXSAPO
DPEQPTDELITTCSSNVSAHADASGRVYILTRDPTPLAANETBAHTVSNAPG
GNIMVAPTLMAAMILMTHFSLILAOELKALDQIIGACACSTLPOIILQHL
GISAFSLHSYRGEINRVAISCLRKLGVPVLRHRRARSRACTLSQGRPALICGYTL
FMAVATTKLKTPIIPASQDLISGFMVAGSGDIYHSVBARPRWFMMLLLLSVGV
GYLLEPNR"

ORIGIN

	mat_peptide	330..902	/product="core protein"
	mat_peptide	903..1478	/product="E1"
	mat_peptide	1479..2756	/product="E2"
	mat_peptide	2757..3407	/product="NS2"
	mat_peptide	3408..5300	/product="NS3"
	mat_peptide	5301..5462	/product="NS4A"
	mat_peptide	5463..6245	/product="NS4B"
	mat_peptide	6246..7586	/product="NS5A"
	mat_peptide	7587..9359	/product="NS5B"
	misc_feature	1479..1559	/note="hypervariable region (HVR)"
	misc_feature	6354..7073	/note="interferon sensitivity determining region (ISDR)"

Query Match 88.3%; Score 1562.6; DB 14; Length 9410;
Best Local Similarity 92.7%; Pred No. 0; Mismatches 129; Indels 0; Gaps 0;
Matches 1640; Conservative 0;

QY	1	CAAAATGACTTCCAGACGCTGACCTCATCGAGGCCAACCTCTGTGGCGGATGATG 60
Db	6981	CATCATGATCTCCCGACAGCCCGACCTCATCGAGGCCAACCTCTGTGGCGGAGATG 7040
QY	61	GGCGGGGAATTAACCGCGCTGAGTGAAGAACAGAGTGAATCTTGAATCTTTCCAG 120
Db	7041	GGCGGGAATATACCGCGCTGAGTGAAGTGAATAGTGTGAATCTTGAATCTTTCCAG 7100
QY	121	CCGCTCGAGCGGAGGATGAGCGGAAGTCCGTCGCGCGGAGATCTTCCGAGAA 180
Db	7101	CCGCTTCGAGCGGAGGATGAGGAGAGTATCCCTTCGCGGAGATCTTCCGAGAA 7160
QY	181	TCCAGAAATTTCCACAGCAGATGCCCGCATGGGCA CGCCGGAATTACAACCTCCGCTG 240
Db	7161	TCCAGAAATTTCCCGCGGCAATGCCCATATGGGAGCGCCGGAATTACAACCTCCAGTG 7220
QY	241	CTGAGACTCTGAGAGCGCCCGGACTACGTCCTCCAGTGTGACATGGGTGCCCATCTGCCA 300
Db	7221	ATAGAGTCTTGAGAGGACCGGACTACGTCCTCCAGTGTGACATGGGTGCCCATCTGCCA 7280
QY	301	CCCTACTAAGACCCCTCTATACCACTCCAGCGAAGAAAGACA GTTGTTCGACAGA 360
Db	7281	CTTACCAAGACCCCTCTCAATACCACTCCAGCGAAGAAAGACG GTTGTCTGACAGAA 7340
QY	361	TCCACCGTCTCTTTCGCTCTGCGGAGCTTGACAAAGCTTTTGATGCTCCGAGCG 420
Db	7341	TCCAGTGTCTCTTTCGCTCTGCGGAGCTTGACAAAGACCTTGCGGAGCTCCGAATG 7400
QY	421	TGGCGCTGACAGCGGACAGCGAACCGCCCTCTGACCAATCTCTCGACGACGCGGA 480
Db	7401	TGGCGCGCGAATGACGCGACCGCCCTCTGACCAAGCCCTCCAGCGACGCGCGAC 7460
QY	481	GCAAGATCGAGGTGAGTGCATCTCTCAATGCCCCCTTGAGGGGAGAGCGGGGAG 540
Db	7461	GCAAGATCGAGGTGAGTGCATCTCTCAATGCCCCCTTGAGGGGAGAGCGGGGAGAC 7520
QY	541	CCCGATCTCAGGACGAGGCTCTTGATCTACCGTGAGTGAAGAGAGCGGAGAGAGCGTCTG 600
Db	7521	CCCGACTCAGGAGCGGCTCTTGATCTACCGTGAGTGAAGAGAGCGGAGAGAGTGTCTG 7580
QY	601	TGCTGCTCATGTCTCTACATGACAGAGCGCTTGATACGCGCATGCGCTCGGAGAGAA 660
Db	7581	TGCTGCTCATGTCTCTACATGACAGAGCGCTTGATACCAATGCGCTCGGAGAGAA 7640
QY	661	AGCAAGCTGCCATCAAGCGGTGAGCAACTTTGCTGTCGTCACCAACAATGCTAC 720

Db 7641 AGCAAGTTACCCATCAATCGTTGAGCAACTCTTTGCTGCGCCACCATATATGCTAT 7700
7721 GTTACCAATCCCGGAGCGGAGCGGCGGAGAAAGATCACTTTGAGACAGATCGAA 780
7701 GCGACAACATCCCGGAGCGGAGCGGAGCGGAGAAAGATCACTTTGAGACAGATCGAA 7760
7781 ATCTGAGAGATCACTACGAGACGCTGCTCAAGAGATGAGGCGAGCGCTCCACAGTT 840
7761 GTCCTGAGAGACCACTACCGGAGCGGAGCGGAGAGATGAGGCGAGCGCTCCACAGTT 7820
841 AAGGCTAAGCTTCTATCACTAAGAGAGAGCTGCGAAGCTGACGCCCCCATTTGCGCAAA 900
7821 AAGGCGAGACTCTTATCCGTAGAGGAGGCGCTTAAGCTGACCGCCCATTTGCGCAAG 7880
901 TCTAATTTGGCTATAGGCGGAGAGAGCGGAGCGGAGAGCTTATCCAGAGAGCGCTTAACAC 960
7881 TCCAAATTTGGCTATAGGCGGAGAGAGCGGAGCGGAGAGCTTATCCAGAGAGCGCTTAACAC 7940
961 ATCCGCTCCGTTGAGAGAGCTTGTGAGAGACACTGAACCACTAATTGACACCACTATC 1020
7941 ATCCGCTCCGTTGAGAGAGCTTGTGAGAGAGACACTGAACCACTAATTGACACCACTATC 8000
1021 ATGCGAAAAATGAGGTTTTCTGCGTCCACAGAGAGAGAGCGCGGAGCGCTGAGCTGCG 1080
8001 ATGCGAAAAATGAGGTTTTCTGCGTCCACAGAGAGAGAGCGCGGAGCGCTGAGCTGCG 8060
1081 CTATCGTTGTTCCAGACTTGGGGGCGTGTGTGCGAGAGAAATGGCCCTCATAGACGG 1140
8061 CTATCGTTGTTCCAGACTTGGGGGCGTGTGTGCGAGAGAAATGGCCCTCATAGACGG 8120
1141 GTCTCACCCCTCCCTCAGGCTGTGATGAGGCTCCTGTCAGATTCAGATTTCTCTGGA 1200
8121 GTCTCACCCCTCCCTCAGGCTGTGATGAGGCTCCTGTCAGATTCAGATTTCTCTGGA 8180
1201 CAGCGGCTGAGATTTCTGCTGAGAACCTGAGAAATCAAGAGAACCCCTATGCGCTTTGCA 1260
8181 CAGCGGCTGAGATTTCTGCTGAGAACCTGAGAAATCAAGAGAACCCCTATGCGCTTTGCA 8240
1261 TATGACACCCGCTGTTTGAACCTCAAGCTGAGTGAAGATCATCCGTTGAGAGAGTCA 1320
8241 TATGACACCCGCTGTTTGAACCTCAAGCTGAGTGAAGATCATCCGTTGAGAGAGTCA 8300
1321 ATTATCAATGTTGACTTGGCCCCCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 1380
8301 ATTATCAATGTTGACTTGGCCCCCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 8360
1381 CGGCTTTATATCGGGGCTCCCTGACTAATTCAAAAAGGCGAGAACTGCGGCTATGCGCG 1440
8361 CGGCTTTATATCGGGGCTCCCTGACTAATTCAAAAAGGCGAGAACTGCGGCTATGCGCG 8420
1441 TSCCGGCGAGGCGGCTGCTGAGCTGAGCTGCGGTAATACCTCAATGTTACTTTGAAG 1500
8421 TSCCGGCGAGGCGGCTGCTGAGCTGAGCTGCGGTAATACCTCAATGTTACTTTGAAG 8480
1501 GCGCTTGAACCTGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1560
8481 GCGCTTGAACCTGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 8540
1561 GCGCTTGAACCTGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1620
8541 GCGCTTGAACCTGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 8600
1621 TTTACGAGGAGGCTATGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1680
8601 TTTACGAGGAGGCTATGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 8660
1681 GACCTGAGGTTGATTAACATCATGCTCTTCAATGTTGCTGAGCGGAGCGGAGCGGAGCT 1740
8661 GACCTGAGGTTGATTAACATCATGCTCTTCAATGTTGCTGAGCGGAGCGGAGCGGAGCT 8720
1741 AAAAGGTTACTACTCAACCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCT 1769

Db 8721 AAAAGGTTACTACTCAACCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCT 8749

RESULT 11
AB049091
LOCUS
DEFINITION
Hepatitis C virus gene for polyprotein, complete cds,
isolate:HCV142.
AB049091
VERSION
AB049091.1 GI:11559448
KEYWORDS
SOURCE
ORGANISM
Hepatitis C virus
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.

REFERENCE
1
Takahashi, K., Iwata, K., Matsumoto, M., Matsumoto, H., Nakao, K.,
Hatahara, T., Ohta, Y., Kanai, K., Maruo, H., Baba, K., Hijioka, M. and
Mishiro, S.
Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients
with hepatocellular carcinoma: the 'progression score' revisited
Hepatol. Res. 20 (2), 161-171 (2001)
11348851
2 (bases 1 to 9547)
Mishiro, S.
Direct Submission
Submitted (19-SEP-2000) Shunji Mishiro, Toshiba General Hospital,
Department of Medical Sciences, 6-3-22 Higashi Oh-i, Shinagawa-ku,
Tokyo 140-8522, Japan (E-mail:shunji.mishiro@po.toshiba.co.jp,
Tel:81-3-3764-8981, Fax:81-3-3764-8992)
location/Qualifiers

FEATURES
source
1..9547
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/isolate="HCV142"
/specific_host="human"
/db_xref="taxon:11103"
/tissue_type="serum"
/note="HCV-genotype: 1b"
1..259
/note="partial"
260..9292
/codon_start=1
/product="polyprotein"
/protein_id="BAB18804.1"
/db_xref="GI:11559449"
/translation="MSTNPKRQRTKXNTRRPODVKFPGGGQVGVGYLLPPRGPPL
GVATYRTTSRQRRQPTPKARQPEGRWAPGYPWPLVGNGLGAWGLSPRG
SRPSPGPTDPRRRRNRLGKVIDLTLCGFADLMGYPLVGAFLGGAALAGVLEAD
GVNATATNLPGCSFSLFLALSLCTIPASAYEVANSGIYHVNDCNSISIVYEAD
VIMHAPGCVPRERNSSRCVVALPTLILARNSSVPTTIRRHVDLVTGTAFCGAY
VGDIQGVFLISOLFTSPRRHETVQSCSLVGHGRHMDMMNSPTALVY
SOLIRPQAVMDVAGAHGVTALATLYSSVGNMAKVLIVMLPAGVDGTHVTGAGV
GSRFLRLTFPTSGASOKIOLINTSGMHNIRNLNCONDLSHTFLALFTYHFNAS
GCEPRLASCRSIDFDQMGPIYAOSSSPORCYCMHVPAPGCIYAPVAGQCPVYC
FTPEPVVGTTRDGRGPTVYGENETVLILNTRPQGMWFGTMMNSVGTCTCGG
PCNIGAGANNLTLCPTDCFRKPEAEATYCGSPMLTPROMVDYPRILMVPCTVNF
TIPKRVYGVGVHRLNAACMKTRGRCNLLEDPRSLSPLLSITTEWOLIPGSEFTLL
PALSTGLIHQNVVDVQVLYGVGSANVSILYIKREYLLPLILADRVCAQIMMLL
IAQEALEMLLVNLNASVAGAHGISLFLVFFCAWYIKSLVGAALVAYVGVPLLL
LILALPRPRAVMDREMAASCGGVFVGLVLTUSPYKVFPLAKLIMVLYQFTTAEAH
LOVWPPLNVRGDAIILTCALHPELIDITGLILALFGLMVLDAQTRVYFVR
AOGILIRACMLVRKAVGAGHYOMAFMKLALITGVVDHLPLRMDAHTGLRLAVYE
PVVSDMETKITITMGATPAACDILISLPUSARGREIILIGAPDSLSCGGMRILAPIT
AYSQTRGLACITITSLTGDKNOVBSEVNSTATOSFLATCNVGCMTYHAGSK
TLAPKPIITOMYTNVDQVLVWQAPRGASLTPCTGSSDLVLTTHADVTYRRRG
DSRGLSLSPRPVSLKSSGGLCPGAVVGLFRAVCTRGAAKVDVPEVEMETL
MSRPFTDNSSPVPAPOTFOVAHLHAPTGSKSKVPAAYAOQYKLVLPVSAATL
GFGAYMSKAGTDNIRGRLITITGAPITYSTYKFLADGCGGAGYDIIICDCHS
TDSYTTIGIGIVDOAETAGARLVLTATPRGSLVTPHPVIEVALSNTEGIPFYCK
ALIPBATKGRHLIFCHSKKKCBELAKLSLSGINAYVTRGLDVSITPISGAVVVA
TDALMTGYTDPSVIDNCVTCTVQTVFSLDPTFTITTTVTPDAVSRQTRGTRGK
RRIYRFTVTRPBRSGMDSVLCECYDAGCAWELTPATRSVLRALVLTNPGLPVCQ

QY 1621 TTCACGGAGCTATGACTAGTACTGTGCCCCCCCCGGGAGACCCGCCCAACAGATATAC 1680
|||||
Db 8531 TTCAGGAGGCTATGACTAGTACTGTGCCCCCCCCGGGAGACCCGCCCAACAGATATAC 8590
|||||
QY 1681 GACCTGAGCTGATTAACATCATGCTCTCCCAATGTGTGCTGCGCAGCATCATCTGGC 1740
|||||
Db 8591 GACTTGAGCTGATTAACATCATGCTCTCCCAATGTGTGCTGCGCAGCATCATCTGGC 8650
|||||
QY 1741 AAAAGGCTACTACTCAGCTCAGCGGTAGCC 1769
|||||
Db 8651 AAAAGGCTACTACTCAGCTCAGCGGTAGCC 8679
|||||
RESULT 12
AF207753
LOCUS AF207753 9374 bp RNA linear VRL 27-APR-2000
DEFINITION Hepatitis C virus strain MD12 complete genome.
ACCESSION AF207753
VERSION AF207753.1 GI:7650223
KEYWORDS
SOURCE
ORGANISM
Hepatitis C virus
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE
AUTHORS Nagayama, K., Kurosaki, M., Enomoto, N., Miyasaka, Y., Marumo, F. and
Sato, C.
TITLE Characteristics of hepatitis C viral genome associated with disease
progression
JOURNAL Unpublished
REFERENCE
AUTHORS Nagayama, K., Kurosaki, M., Enomoto, N., Miyasaka, Y., Izumi, N. and
Sato, C.
TITLE Direct Subinjection
JOURNAL Submitted (23-NOV-1999) Second Department of Internal Medicine,
Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-ku,
Tokyo 113-8519, Japan
FEATURES
source
1..9374
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/strain="MD12"
/db_xref="taxon:11103"
/note="Isolated from a patient with persistently-normal
ALT values (ASC-2)"
330..9362
/codon_start=1
/product="polyprotein"
/protein_id="AAF65943.1"
/db_xref="GI:7650224"
/translation="MSTNPKPQKTKRNTNRPRPOVKKPRGGQIVGVLLPRLPRRL
GYRATRTKTSEROPFRGRPIFKVROBRTVAQPEYFPLVIGMGMGWALPSPRG
SRPSWGPDPRRRRRNLIKVIDLTCGFDLMGVIPLVGAAPLGAARVLAHGVLEP
GVNATGNLPGCSFSLIALSLCLTPASAEVNRVSGMYHVTNDCNSISVYAAD
MIMHAPGCVPCRENNSRCAVALPTTLARNAVSPTTIRHVDLLVGAAPFCVAM
VDDLSGVFLISQLFTFSPRHRHTVDDNCISYIPGHSRHMAMMMNMWSTLIV
SGLRTPQAVLVVAGAHGWVLAGLAAYISAGNMAKVLVLLPAGVNGQRTISGTO
GHTTGLVLFSSGSQNTQLINTNGSWHINTALNCNDLSTGLFALVYHKNAS
GCPARVACRSIDTEPAQWGPITVAQPSIDORPCMHAPRCQIVPASOVCGVVC
FTSPVVGITDRSDPGVTSWGENETDVLINNTRPPOGNMGCTMMSTGCTKCG
PCNIGGVGNNTLTCTPCDFCRKPEATYKCGSPMLPRCLVDVPRYLMHPCVNF
SIFKRVYGVGEHRLNAACNMTGRGCGLEDNRSELSPLLSTTBQVLPSPCTTL
PALSTGLIHLHONVVDVLYIGISAVNSFYVKMDYVILLPLLDARCACTMMML
TQAEALENLVVLANASVAGAHGLPLVFFCAWYIKGRVFGAALVYVWPLL
LILALPRAVYAMDREMAASCGAVLVGLMILTLSPYKFLARLIMWLYLTRFAH
LOVWVPLNVRGRDAIPLTCATHPELFDITKLILALIGPLMVLQAGITVPEVR
AOGILRACMLVRKAGAHYVOMAEMKALATGTYYDHLTIRGNAHTGRDLAAVE
PVFSPMTEKIIITWGADTAACGDIILGLVSARGRLEILGPADELBOGMRLLAPIT
AVSOQTRGLGCIITSLTGRDKQVGEVYVSTATQFTLATCNAGVCTVYHAGSK
TLAEGKGLITQMTINVDQDLVGMQAPGARSMPTCTGSSDLYLTRADVLPVRRG
DGRGLSPRVSYLKGSGGGLCPSGHVGIFRAAVCTRVAAVDFVPEVSEMT
MSPVFTDMSPPAVPQTFQVHLHAPLTSKSTKVPAAVYAAGYKVLVNLNPSVAATL

ORIGIN

GFAYMSKAVGTDENIRGTVAITTTGAPITVITYGKFLADGCGGAVDIIICDECHS
TDSYTIILGIVLLOAETAGARLVLAATATPGSVATVPHNIEVVALSNTEIPFYGK
AIPIETIKGRIHLIFGSKKKDELAALKSLGAINAAVYPIGELDVYVLTGGVVA
TDALMTGTPGFDSDIDNCVOTQVDPTEPTITDTPVPOAVRTOGRGTGG
RRGIYRPTTPBGRSMPFSDSVLECDYDAGCAMVELTPAETSVRTLPAVIRPGIPVOC
DHLFMEBVFGLTRIDHAPLSOTQKGDNPVNAIOATVCARQAPPSMDMMKC
LTRLKPTLHGPTPLVIRLGAOVNVTLLPITKXIMTCMSADLEVSTVWLVGSLA
ALAAVCLITGSSVIVGR11LISGRPA11IDREVLVREBDEMECSQIPLYEQMOLAE
OFKOKAGLODTATKOAAAPVYESKMAALAEFMWGMNMFISGIQYLAGSLTLPN
PAISLMAFTLSTVSPILTOSTLTFNLIGVNAOOLPPSAASAFVAGL1AGAVLS1
GLGVYVDIILAGYAGVAGALVAKWMSGEBPSTEDVNLILPALSICLVYGVCA
ILKRVHGBEGAVQVMNRLIFASRGHVSFTIRVPSDAPARVTQILSNLTITQLR
RLHOMINEDCSTPCSGSMLRDVMDI1CVLSDFTFWLSQKLLPRLPGVPLSCGRYK
GVMRGDIGIMOTTCPGCAOITGVHNGSMR1VGPCTGNTMHTGPPINAYTTPCTPSP
APNYSBALMRYAAEEVYEVTEGDFPHYVGTNTDNCPCCPVAPPEFTEBQVRLR
YAPACKPLIREHYFOVLNOYLVGSQLPCGPEPDVAVLSMLTDPHIT1ETAKRL
AKGSPPLSSASASQLSAPSLKATCTTHKSPDDALIBANILRQENGNITRPSER
KVVLIDSEDP1LRABEDEREISVPAEL1LTKRKPFPALP1WARPDPYNPLESWKDPY
VP1PVHGGPLPPTKGPPLP1P1RRRTVTL1ESTVSLAEATATFGSSGASDST
ATAPDDQTSDDGDSDVESYSMPLEGEPPDLDSDGMSVSEAGEBVDVCCMS
YTMGALITPCAAEESKLPIVALNSLLRHNNMYATTSRACOROKKVP1DRLOVD
DHYRVDV1KEMKAKASTVYKALLSVBEAC1TPHSASKRGYAKVDVNR1SSKVNHI
RSWKDLIEDTEPTINTTWAKNBVFCTQPEKGRKARLIVPDLGVRCERKALVD
VSTL1PQAVWSSYGFQYSPQARVEPLVNTWKSCKCPMGEYDTRCDSTVTEIDIRV
BES1YQCCDLA1PEAROV1RSLTERLVYGGPLTNSKQNCGRRCRAGSVLT1SCGNTL
TCYLKASACRAKALODCTMLVNDGDLV1CESAGTODASLAEVTEAMTRVYAPG
DPPOPEYDELEITSCSSNVSAVHADSGRYYL1FRDPTPLPARAMEATRT1PNSML
GNITMVAPTLMABM1LMTWPS1LLAEOLEKALD1COIYGCVS1EPLD1POLIERLH
GLSAPLSHSYSPGEINRYASCTRLKCTPPLRVHHRARSYRAKL1SGCGRATVSGKL
FNMAVR1KL1P1PAASRLD1SGM1FVAG1SGD1YHLS1RARR1PML1LL1SVG
GIVLL1P1NR"

Query Match 88.0%; Score 1557.8; DB 14; Length 9374;
Best Local Similarity 92.5%; Pred. No. 0; Gaps 0;

Matches 1637; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

QY 1 CAAATGACTTCCAGAGCGTGAAGCTCATCAGAGCAACTCTGTGCGGCGATGAGATG 60
Db 6981 CACCATGACTCCCGGAGCGCTGAGCTCATTCAGAGCAACTCTGTGCGGCGAGAGATG 7040
|||||
QY 61 GCGCGGAGCATTTACCCGCGTGAAGTCAGAGAAAGGTAATCTCTGACTCTTTGCAC 120
Db 7041 GCGCGGAGCATTTACCCGCGTGAAGTCAGAGAAAGGTAATCTCTGACTCTTTGCAT 7100
|||||
QY 121 CCGCTCCGAGGAGAGAGATGACGCGGAATGTCCTCCGCGGAGAGATCTCGCGAA 180
Db 7101 CCGCTCCGAGGAGAGAGATGACGCGGAATGTCCTCCGCGGAGAGATCTCGCGAA 7160
|||||
QY 181 TCACAGAAATTCACACAGCATGCCGATGAGGACGCCGGAATTAACACCTTCGCTG 240
Db 7161 CCCAGAGATTCCTCCCGACGCTTGCCCATATGAGGACGCCGGAATTAACACCTTCGCTG 7220
|||||
QY 241 CTGAGTCTGGAAGAGCCCGGAGACTACGTCCTCCAGTGTAATCGAGTCCCACTGCA 300
Db 7221 CTGAGTCTGGAAGAGCCCGGAGACTACGTCCTCCAGTGTAATCGAGTCCCACTGCA 7280
|||||
QY 301 CCTACTAGACCCCTCTCTATACCACTTCCAGGAGAAAGAGACAGTTGTTCTGACGAA 360
Db 7281 CCTACTAGACCCCTCTCTATACCACTTCCAGGAGAAAGAGACAGTTGTTCTGACGAA 7340
|||||
QY 361 TCACCGGTGCTCTTGCCTGCGGAGGCTTCCCAAAAGCTTTGGTAGCTCCGAGACG 420
Db 7341 TCACCGGTGCTCTTGCCTGCGGAGGCTTCCCAAAAGCTTTGGTAGCTCCGAGACG 7400
|||||
QY 421 TCGGCGGTGACAGCGGAGCGGCAACCGCCCTCTCTGCAATCTCTCGAGCAGCGGCA 480
Db 7401 TCGGCGGTGACAGCGGAGCGGCAACCGCCCTCTCTGCAATCTCTCGAGCAGCGGCA 7460
|||||
QY 481 GCAAGATCTGACGTTGAATCGTATCTCTCATTGCCCCCTTGAAGGGGAGACCGGGGAC 540
Db 7461 GCAAGATCTGACGTTGAATCGTATCTCTCATTGCCCCCTTGAAGGGGAGACCGGGGAC 7520
|||||


```

QY 541 CCCGATCTCAGCGAGGGGCTTGCTACCGTGAGTGAAGGCGGGTGAGACGTGCTC 600
DB 7521 CCGGATCTCAGCGATGGGCTTGCTACCGTGAGTGAAGTGGTGAGACGTGCTC 7580
QY 601 TGTGCTGATGTCCTACACATGAGAGGCGCTCTGATCAAGCCATGCGCTGCGAGAA 660
DB 7581 TGTGCTGATGTCCTACACATGAGAGGCGCTCTGATCAAGCCATGCGCTGCGAGAA 7640
QY 661 AGCAAGCTCCCATCAACCGCTTGAGCAACTCTTTGCTGCTGATCAACAATGCTTAC 720
DB 7641 AGCAAGCTCCCATCAACCGCTTGAGCAACTCTTTGCTGCTGATCAACAATGCTTAC 7700
QY 721 GGTACCAATCCCGGAGCGGAGCGGAGCGGAGAGGTCACCTTTGACATGTCGA 780
DB 7701 GGTACCAATCCCGGAGCGGAGCGGAGCGGAGAGGTCACCTTTGACATGTCGA 7760
QY 781 ATCTGAGAGATCACTACAGAGAGCGTGTCAAGAGATGAGAGGCGGAGCGTCAAGTT 840
DB 7761 GTCCTGAGAGACCACTACCGGAGCGTGTCAAGAGATGAGAGGCGGAGCGTCAAGTT 7820
QY 841 AAGGCTAAGCTTCTATCAGTAGAGAGAGCGCTGCAAGTGAAGCCGCCACATTGCGCA 900
DB 7821 AAGGCTAAGCTTCTATCAGTAGAGAGAGCGCTGCAAGTGAAGCCGCCACATTGCGCA 7880
QY 901 TCTAATTTGGCTATGGGCGAAGAGAGCGTCCGAGCTATCCAGAGGCTTAAACAC 960
DB 7881 TCTAATTTGGCTATGGGCGAAGAGAGCGTCCGAGCTATCCAGAGGCTTAAACAC 7940
QY 961 ATCTGCTCCGTGTGAGAGAGACTTGTGAGAGACCTGAAACACCAATTGACACACCATC 1020
DB 7941 ATCTGCTCCGTGTGAGAGAGACTTGTGAGAGACCTGAGAGACCTGAAACACCAATTGAC 8000
QY 1021 ATGGCAAAAATAGAGTTTCTGCTGCAACCAAGAGAGAGCGGCAAGCCAGCTCGC 1080
DB 8001 ATGGCAAAAATAGAGTTTCTGCTGCAACCAAGAGAGAGCGGCAAGCCAGCTCGC 8060
QY 1081 CTATGCTGTCCAGACTTGGGGGTCCGTGTGAGAGAAATGGCCCTATGAGCGT 1140
DB 8061 CTATGCTGTCCAGACTTGGGGGTCCGTGTGAGAGAAATGGCCCTTATGAGCGT 8120
QY 1141 GTCTCAGACCTCCCTCAGGCTGATGAGGCTCCTCTGATCCAGATTCAGATTCCTCTG 1200
DB 8121 GTCTCAGACCTCCCTCAGGCTGATGAGGCTCCTCTGATCCAGATTCAGATTCCTCTG 8180
QY 1201 CAGCGGCTGAGTCTCTGTGAAAGCGCTGAAATCAAGAGAGACCCCTATGGCTTGA 1260
DB 8181 CAGCGGCTGAGTCTCTGTGAAATCAAGAGAGACCCCTATGGCTTGA 8240
QY 1261 TATGACACCCGCTTTTGAATCAACAGTCACTGAGAAATGACATCCGTGTGAGAGTCA 1320
DB 8241 TATGACACCCGCTTTTGAATCAACAGTCACTGAGAAATGACATCCGTGTGAGAGTCA 8300
QY 1321 ATTATCAATGTTGATCTTGAGCCGCCAGAGCAGACAGGCTAAGTGTGCTCAGAG 1380
DB 8301 ATTATCAATGTTGATCTTGAGCCGCCAGAGCTAAGTGTGCTCAGAG 8360
QY 1381 CGGCTTATATGAGGAGGCTCCCTGACTAATTTCAAAAGGAGAACTGCGGCTATCCCG 1440
DB 8361 CGGCTTATATGAGGAGGCTCCCTGACTAATTTCAAAAGGAGAACTGCGGCTATCCCG 8420
QY 1441 TGGCGGAGAGGCGGCTGAGTGAAGTGTGAGTGAATGATCCCTCAGATGTTACTTGA 1500
DB 8421 TGGCGGAGAGGCGGCTGAGTGAAGTGTGAGTGAATGATCCCTCAGATGTTACTTGA 8480
QY 1501 GCGCTGAGAGCTGTGAGCTGCAAGCTCAGAGCTGACGATGCTGTGTGCGAGAG 1560
DB 8481 GCGCTGAGAGCTGTGAGCTGCAAGCTCAGAGCTGACGATGCTGTGTGCGAGAG 8540
QY 1561 GCGCTGTGTGTTATCTGTGAGAGCGGAGAACCAAGAGAGAGCGGCGAGCTTACAGATC 1620
DB 8541 GCGCTGTGTGTTATCTGTGAGAGCGGAGAACCAAGAGAGAGCGGCGAGCTTACAGATC 8600
QY 1621 TTCAGGAGAGCTATGACTAGTACTGCGCCCGCGGAGAGCGGCGCAACAGAAATAC 1680

```

```

DB 8601 TTCAGGAGAGCTATGACTAGTACTGCGCCCGCGGAGAGCGGCGCAACAGAAATAC 8660
QY 1681 GACTGAGTGTATTAATCATCATGCTCTCAATGATGTGCGGCGAGATGATCTGCG 1740
DB 8661 GACTGAGTGTATTAATCATCATGCTCTCAATGATGTGCGGCGAGATGATCTGCG 8720
QY 1741 AAAAGGTATTAATCACTCAACCGGTGACC 1769
DB 8721 AAGAGGTGTACTCACTCAACCGGTGACC 8749

RESULT 13
HPCGENANTI
LOCUS HPCGENANTI 9425 bp mRNA linear VRL 14-JUL-1994
DEFINITION Hepatitis C virus Taiwan mRNA, 5' end.
ACCESSION M84754
VERSION M84754.1 GI:329763
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepadnaviruses.
REFERENCE 1 (bases 1 to 9425)
AUTHORS Chen, P. J., Lin, M. H., Tai, K. P., Liu, P. C., Lin, C. J., and Chen, D. S.
TITLE The Taiwanese hepatitis C virus genome: sequence determination and
mapping the 5' terminus of viral genomic and antigenomic RNA
JOURNAL Virology 188 (1), 102-113 (1992)
MEDLINE 92230206
PUBMED 1314449
COMMENT Original source text: Hepatitis C virus (strain Taiwan) cDNA to
mRNA.

FEATURES
source location/Qualifiers
1..9425 /organism="Hepatitis C virus"
/mol_type="mRNA"
/db_xref="taxon:11103"

ORIGIN
Query Match 88.0%; Score 1557; DB 14; Length 9425;
Best Local Similarity 92.6%; Pred. No. 0;
Matches 1635; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 5 ATGACTTCCCAAGACGTTGACTGATGAGGCCAAGCTCTGTGCGGCGATGAGATGGCG 64
DB 6997 ATGACCCCAAGACGTTGACTGATGAGGCCAAGCTCTGTGCGGCGATGAGATGGCG 7056
QY 65 GGGACATTACCGCGGTGAGTGAAGAAAGTATGTAATCCGTGAGCTTTGACCCGC 124
DB 7057 GAAACATTACCGCGGTGAGTGAAGAAATGATATCTTGACCTTTGACCCGC 7116
QY 125 TCCGAGCGAGAGATGAGCGGAAAGTGTCCGTCCCGCGAGATCTCGGAAATGCA 184
DB 7117 TCCGAGCGAGAGATGAGCGGAAAGTGTCCGTCCCGCGAGATCTCGGAAATGCA 7176
QY 185 AGAAATTTCCACAGAGATGCGCGCATGGGCAAGCCCGATTAACAACCTCCGCTGCG 244
DB 7177 GGAAGTTCCCTCAAGGTGTCCCGATGAGCAAGCCCGATTAACAACCTCCGCTGCG 7236
QY 245 AGTCCGGAAGGCGCGGACTAGTCCCTCCAGTGTGATCATGGGTGCCACTGCGACCTA 304
DB 7237 AGCTCGGAAGGCGCGGACTAGTCCCTCCAGTGTGATCATGGGTGCCACTGCGACCTA 7296
QY 305 CTAAAGACCCCTCTATACCACTTCAACGAGAAAGAGCAAGTTGTTCTGACAGATGCA 364
DB 7297 TCAAGGCGCTCCATATCACTTCAACGAGAAAGAGAGCGGTTGCTGACAGATGCA 7356
QY 365 CCGTGTCTTCTGCGCGGAGCTTGCACAAAGCTTTTGTGATGCTCCGACCTGCG 424
DB 7357 CCGTGTCTTCTGCGCGGAGCTTGCACAAAGCTTTGATGATGCTCCGACCTGCG 7416
QY 425 CCGTGAAGAGGAGCAAGCGCCCTCTGACCAATCTTCGAGAGCGGAGAGAG 484

```

Db 7417 CGCTGTGATCGGCAAGGCAAGCCCTCTGATCAGCCCTCCGACGACGCGAG 7476
 Qy 485 GATCTGACGTTGAGTGTCTCTGATGACGCCCCCTTGAGGGGGAGCGGGGACCCCG 544
 Db 7477 GATTCGACGTTGATGATGCTCTCTCTATGCCCCCTTGAGGGGGAGCGGGGACCCCG 7536
 Qy 545 ATCTGACGACGGGCTTGTGTCTACCGTGAAGTGAAGAGGCGGTGAAGACGTCTGTGCT 604
 Db 7537 ATCTGACGACGGGCTTGTGTCTACCGTGAAGTGAAGAGGCGGTGAAGACGTCTGTGCT 7596
 Qy 605 GTTCGATGCTCTACATGAGACAGGGGCTCTGATCAGCCGATGCGGTGGAGAGAAAGCA 664
 Db 7597 GCTCGATGCTCTACATGAGACAGGGGCTCTGATCAGCCGATGCGGTGGAGAGAAAGCA 7656
 Qy 665 AGCTGCCCATCAACGCGTTGAGCAACTCTTGTGCTGCTCACCACCAATGATGTCTAGCTA 724
 Db 7657 AGCTGCCCATCAATGCGTTGAGCAACTCTTGTGCTGCTCACCACCAATGATGTCTAGCTA 7716
 Qy 725 CCACATCCCGACGCGACGCGGACGCGGACGAGAAAGGTCACTTTGACAGACTGCAAAATCC 784
 Db 7717 CCACATCCCGACGCGACGCGGACGCGGACGAGAAAGGTCACTTTGACAGACTGCAAGTCC 7776
 Qy 785 TGGAGAGTACCTACACGAGACGCTGCTCAAGAGATGAAGGCGGAGGCTCCACAGTTAAG 844
 Db 7777 TGGAGAGACCTACACGAGACGCTGCTCAAGAGATGAAGGCGGAGGCTCCACAGTTAAG 7836
 Qy 845 CTAACTCTCTACAGTGAAGAGAGCTGCAAGCTGACGCGCCCACTTGGCGCAATCTA 904
 Db 7837 CTAACTCTCTATCTGTAGAGAGAGCTGCAAGCTGACGCGCCCACTTGGCGCAAGTCTA 7896
 Qy 905 AATTGGCTATGAGGCAAGAGAGCTGCGGAACCTATCCAGCAAGGCAATTAAACAATCC 964
 Db 7897 AATTGGCTATGAGGCAAGAGAGCTGCGGAACCTATCCAGCAAGGCAATTAAACAATCC 7956
 Qy 965 GCTCCGTGAGAGAGACTTGTGGAAGACACTGAACACCAATTGACACACCATCATAG 1024
 Db 7957 GCTCCGTGAGAGAGACTTGTGGAAGACACTGAACACCAATTGACACACCATCATAG 8016
 Qy 1025 CAAAAAATAGAGTTTCTGCTGCTCAACGAGAGAGAGGCGGCAAGCAGCTCGGCTTA 1084
 Db 8017 CAAAAAATAGAGTTTCTGCTGCTCAACGAGAGAGAGGCGGCAAGCAGCTCGGCTTA 8076
 Qy 1085 TGTGTTTCCAGACTTTGGGGGCTCGGTGTGCGAGAAATGGCCCTTAATGACGTGCT 1144
 Db 8077 TGTGTTTCCAGACTTTGGGGGCTCGGTGTGCGAGAAATGGCCCTTAATGATGTGCT 8136
 Qy 1145 CCAACCTCCTCAGAGCTGATGAGGGGCTCTGCTGATCGAATTCAGATATCTCTGAGACGC 1204
 Db 8137 CCAACCTCCTCAGAGCTGATGAGGGGCTCTGCTGATCGAATTCAGATATCTCTGAGACGC 8196
 Qy 1205 GGGTCAGATTCCTGAGTGAAGCGCTGGAATCAAGAGAGACCCCTAATGGCTTTGATATG 1264
 Db 8197 GGGTCAGATTCCTGAGTGAATTCCTGGAATCAAGAGAGAGCCCTAATGGCTTTGATATG 8256
 Qy 1265 AACACCCGCTGTTTGACTCAAGACTCACTGAGATGACATCCGTGTAGAGAGTCAATT 1324
 Db 8257 AACACCCGCTGTTTGACTCAAGACTCACTGAGATGACATCCGTGTGAGAGTCAATT 8316
 Qy 1325 ATCAATGTTGATCACTTGGGCCCCGGAAGCGACAGAGCCCTAAGTGTCTCAAGAGCGGC 1384
 Db 8317 ATCAATGTTGATCACTTGGGCCCCGGAAGCGACAGAGCCCTAAGTGTCTCAAGAGCGGC 8376
 Qy 1385 TTTATATCGGGGGTCCCTGATCAATTTCAAAAGGGGAGACTGCGGCTATCGCCGATGCG 1444
 Db 8377 TTTATATCGGGGGTCCCTGATCAATTTCAAAAGGGGAGACTGCGGCTATCGCCGATGCG 8436
 Qy 1445 GCGGACGCGGCTGCTGACGACTAGCTGCGGTATACCTCTCAATGTTACTTTGAAAGGCT 1504
 Db 8437 GCGGACGCGGCTGCTGACGACTAGCTGCGGTATACCTCTCAATGTTACTTTGAAAGGCT 8496
 Qy 1505 CTGACGCTGTGAGAGTGAAGGCTCAAGACTGACAGATGCTCGTGTGCGAGACGCGC 1564
 Db 8497 CTGACGCTGTGAGAGTGAAGGCTCAAGACTGACAGATGCTCGTGTGCGAGACGAGC 8556

Qy 1565 TTGTCTATCTGTGAGAGCGCGGGAGCCAGAGAGAGCGCGGAGCTTACAGAGTCTTCA 1624
 Db 8557 TTGTCTATCTGTGAGAGCGCGGGAGCCAGAGAGAGCGCGGAGCTTACAGAGTCTTCA 8616
 Qy 1625 CGAGGCTATGATCAAGTACTTGTGCCCCCGCGGGAGCCGCGCCCAACAGAAATGAGCC 1684
 Db 8617 CGAGGCTATGATCAAGTACTTGTGCCCCCGCGGGAGCTTCCCAACAGAAATGAGACC 8676
 Qy 1685 TGGAGTTATATCAATCAATGCTCTCCCAATGTGTGCTGCGGACAGATGATCTGGCAAAA 1744
 Db 8677 AGAGTTATATCAATCAATGCTCTCCCAATGTGTGCTGCGGACAGATGATCTGGCAAAA 8736
 Qy 1745 GGGTACTACTACCTCAACCCGAGACC 1769
 Db 8737 GGGTACTACTACTACCCGAGACC 8761
 RESULT 14
 AB049090
 LOCUS
 DEFINITION
 Hepatitis C virus gene for polyprotein, complete cds,
 ACCESSION
 AB049090
 VERSION
 AB049090.1 GI:11559446
 KEYWORDS
 SOURCE
 ORGANISM
 Hepatitis C virus
 Hepatitis C virus
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
 REFERENCE
 1
 AUTHORS
 Takahashi,K., Iwata,K., Matsumoto,M., Matsumoto,H., Nakao,K., Harahara,T., Ohta,Y., Kanai,K., Maruo,H., Baba,K., Hijikata,M. and Mishiho,S.
 TITLE
 Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients with hepatocellular carcinoma: the 'progression score' revisited
 JOURNAL
 HEPATOLOGY
 PUBMED
 11348851
 REFERENCE
 2 (bases 1 to 9573)
 AUTHORS
 Mishiho,S.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (19-SEP-2000) Shunji Mishiho, Toshiba General Hospital, Department of Medical Sciences; 6-3-22 Higashi Oh-1, Shinagawa-ku, Tokyo 140-8522, Japan (E-mail:shunji.mishiho@po.toshiba.co.jp, Tel:81-3-3764-8981, Fax:81-3-3764-8992)
 FEATURES
 source
 location/Qualifiers
 1..9573
 /organism="Hepatitis C virus"
 /mol_type="genomic RNA"
 /isolate="HCV140"
 /specific_host="human"
 /db_xref="taxon:11103"
 /tissue_type="serum"
 /note="HCV-genotype: 1b"
 <1..310
 /note="partial"
 311..9343
 /codon_start=1
 /product="polyprotein"
 /protein_id="BAB18803.1"
 /db_xref="gi:11559447"
 /translation="MSTNPKPQRTKENTNRPODVKPGGQIVGGVLLPRRGRPLGVATRTKRSRSGRROPKROEGRAMAGPGVWPLVGNELGMAVLLSPRGSRPWSGPDPRRSRNGKVDLTLCGPADLMGVIPLVGAFLGAARALAGVLEDCGNVATGRLPCCPSIFLLALSLTTPASAEVRANSGVYHVTNDCNSIYVEADMIHMPGCVPCVRDENVSRVVALPTLAAANSIPFTTIRHVDLVGAALCSAMTVGDICGSVFLVSQLFPSPRRYETVQNCISLYGHVSGHRMADMMNMSPTALVVSQALRIPOAVVDVAGAHMVGVLAYLSMGNMAKYLIVMLFAGVDNHTVGGQAGOHARFSLSPSSASOKIOLINTNGSMHIRTALNCNDSLHTFIALFYAHKFNASGCPERMASCPRIKPDQGWGPISHVNPISDORPCYHVARPGGIYVPSAGCPVYCFPSPPVVGTTDRFGVFTYTGNEETVLLNTRPPOGMFGCTMMNGTFTTCGGPPCNIGAGNNTLTCPTDCPKRHEATVYTCGSGPWLPRCIYDVPYRLMPCVTPNRTIFKRMVVGVEHRLNAACNMTGRCNLEDRSRSLSPLLSTTEWQIMPCSTTL

PALSTGLIHLHONIIVDOYLYGVGSAVSVISVIMKEVYVILLFLLADARVACILMMLL
 IAAEBALENUVYLANASVAGAGTISTLVEFCAANTIKKIVPGAAVAYFVWPLLL
 LIALPREAYAMDEMAASCGVTFVGLVLLISPHKFLALIMLOFTIRACAC
 LQVMIPLNVRGGRDAIILLTCVHPELVFEIKLILALIGPLMVLOAGLTRVYVR
 AOGILIRACMLVRVAGGVYVOMALMKLALGTGVYDHLPLMDMAAGLRLVAVE
 PVIFSMETKILITMGADTAAAGDIIILGPASARGREIILGPDSEGEQMRDLAPLT
 AVSOGIRGLIGCIITSLTGRDKQVEGVENVSTATOCISFLATCVNCGMCHVYASAK
 TLABKPIITOMVYNDODLVGMAPRGARSILPCICGSGLDLYVRHADVIRERGC
 DERGSLSRPISLYKSGSGPLICSGHNAVIGIRAVCTRGAKADVIRVSMET
 MRSVPTDNSSPPAVPQTFQVAHIIHAPSGSKSTVPAALAAAGVLYLVNPSVATL
 GGAVMSKAGVDPNIRITGVRTITGPAITVSTYKGLVADGGSGAGVILIDEGHS
 TOSTILIGITVLOAETAGARLVATAPPGSVYVPHNIEVLSNTEIPFGK
 ALPLEAIIKGRHLIFCHSKKODELAAKISGLVNAVAYYRGDADVIRPISGVVVA
 TALMNGYTGDPDSVITDCTCTVQYVDFSLDPFTTETTTVPDAYSRSORRTGRG
 RRGTYRFTVPERSGMLDSVLCEDTDGCAVTELPATVSRILALIPGLPYCQ
 DILFEPESVFTGLTHIDAFHLSQTKOAGNPPYLVAVOATCARAOPPSMOQMKC
 LILRKPTLHGPPPLVRLGAVONEVTLTHITKYIMACMSADLEVTSITVLGVGLA
 ALAAVCLTTGSSVIVGRILISGKPAIIPDEVLVQFEDEMEECASHLYIEQOMOLAE
 OFOKAKGLLOTATKOAEAPAVVESKMALETPEMAKHMNPIISGIOYLAGSTLBN
 PALASIMATASITSPLOTHTLLFNILGEMVAQOLAPPSASAIPGAGIACAUVSI
 GAGKLVLDIAGVAGVAGLVAFKIMSGESTEDLVNLLPAILSPGALVGVVCA
 ILIRHVGPGEAVQOMNRILAFASRGNHSPHYVPESDAAVTOILSLITLOLK
 RLHOWINBEDCTPCSGSWLSDVMDICTVLVDFKTWLOSLPRLGVPFLSGORGYK
 GVRGDIQMOTPCGAOIAGHVNGSMRIVGPRCSNTHGEPINAYATGTCTSP
 APVSRALMRVAEEVEITRVGDPHYVGMCTTDNIKCPQVAPAPPEFTELDVRLHR
 YAPACKPLRDEYTFOVGNQVYVSQOLPCEPEPDVYVLSMLTDSHITAEAKRL
 AAGSPSLASSASQSLASBLKATCTTRDOSPADLILEANLRQMGAGITRVESEN
 KVIILDSFEPRLABEDEREVSVAELIRKSRKPPALPVARMDWIPALLESMDKY
 VPPVHVCPLPTKAPPIPPPRKRTVULTESTVSALAEATKTESGSSSVADSGT
 ATPPOPSDDGDAGSDVESYSMPLEGGPDDLSDEMSITVSEASDVSIGSCMS
 YMTGALITPCAEBSKLPINLNSLILHNMVNAVATKRSASOROKRATFDLOLD
 DHYRDLKEMKAKASTYKAKULSVEBCKLTPHRSARSKRGYAKRNLRSOAVNI
 KSVWKULBOTOPIVITVMAKNEVFCVPERKGRPARLIVPDLGVRCEKMAVLI
 VVSTLQAVMGAASYFOYSPGQREFLVNAWSKRCMPGVARFDFSVIESTDIRV
 EBSIYOCCLAPARQAIRSLTERLYVGGDLVNSKQNCYRCRASGVLTJSCGTL
 TCYLKATACRAKLODCTMLVCGDVLVYVCSAGTOEDPAARVLEAMETARSAPG
 DPPOPEYDELITSCSSNVSVAHSDAGSKVYVLTDPDPLAASAEETAHPVNSWL
 GNIINAPLTMARMIIMTHPFSILAOBOLKALDCOITGATYSIEPDLPIQLOH
 GJASFSLSHSYSGEINRVASCKLGLVPLRVHRHARSRAULISQGGGAACGTYL
 FMAVATVTKLKTPIIPASOLDLSGWFIAVSGGDIYHSLSRARPRFMFCLLISGV
 GYLLPNR"
 9344..9396
 /note="upstream of poly-U stretch"
 9397..9439
 /note="poly(U) length is indistinct
 poly(U) stretch"
 9440..9573
 /note="3' X tail"

ORIGIN

Query Match 88.0%; Score 1557; DB 14; Length 9573;
 Best Local Similarity 92.6%; Pred. No. 0;
 Matches 1635; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

3'UTR
 3'UTR
 3'UTR

5 ATGACTCCCAAGCGCTGACCTCATGAGGCCAACTCCCTGTCGCGGCAATGAAATGGCG 64
 6966 AAGACTCCCAAGCGCTGACCTCATGAGGCCAACTCCCTGTCGCGGCAATGAAATGGCG 7025
 65 GGAACATTACCCGCTGAGTCAAGAAACAAGGTAGTAATCTGAGACTCTTTGACCCGC 124
 7026 GGAACATTACCCGCTGAGTCAAGAAACAAGGTAGTAATCTGAGACTCTTTGAAACGC 7085
 125 TCCGAGCGGAGAGATGAGCGGAAAGTGTCCGTCCCGCGGAGATCTTCGCGAAATCCA 184
 7086 TTCGCGCGGAGAGATGAGCGGAAAGTATCCGTTCCGCGGAGATCTTCGCGAAATCCA 7145
 185 AGAATTTCCACACGAGGATGCCCGCANTGGSCAGCCCGGATTTCACACCTCCGCTGCTG 244
 7146 GGAAGTTCCCGCGGCGCTGCCGTATGGSCACGCCCGGATTACAAACCTCCCACTATAG 7205
 245 AGTCTTGAAGGCCCCGAGCTACGTCCTTCAGTGTATACATGGGTGCCCACTGCCACTTA 304
 7206 AGTCTTGAAGGACCAAGACTACGTCCTTCGAGTGTATACATGGGTGCCCACTTGCACCTTA 7265

305 CTAAGACCCCTCCTATACCACTCCACGAGAAAGAGACAGTTGTTCTGACAGATCCA 364
 7266 CTAAGGCCCTCCAAATACCACTCCACGAGAAAGAGACAGTTGTTCTGACAGATCCA 7325
 365 CGGTGCTTCTGCGCCTGGCGAGACTTGGCCAAAGGCTTTTGGTAGCTCCGACCGTGG 424
 7326 CGGTGCTTCTGCGCCTGGCGAGCTTGGTCAAAAGACTTTGGACACTCCGAGTCTGG 7385
 425 CCGTGACAGCGGACAGGCAACCGCCCTCTTGACCAATCTTCGACGACGCGGAGAG 484
 7386 CCGTGACAGCGGACAGGCAACCGCCCTCTTGATACGCCCTCCATGACGAGAGAGAG 7445
 485 GATCTGACGTTGAGTGTATCTCTCATGATCCCCCTTGAAGGGGAGCCGGGGAGACCC 544
 7446 GATCTGACGTTGAGTGTATCTCTCATGATCCCCCTTGAAGGGGAGCCGGGGAGACCC 7505
 545 ATCTGACGCGAGGCTTGTGATCTACCGGTAGTGAAGAGCCGGTGAAGAGCTGCTGCT 604
 7506 ATCTGACGCGAGGCTTGTGATCTACCGGTAGTGAAGAGCTGATGAGATGCTGCTGTT 7565
 605 GCTGATGCTCTACATGATGAACAGGCGCTGTATCAACGCAATGCGCTGCGAGAAAGA 664
 7566 GCTGATGCTCTACATGATGAACAGGCGCGCTGTATCAACCAATGCGCTGCGAGAAAGA 7625
 665 AGCTGCCATCAACAGCGCTTGAGCAATCTTTGCTGCTGATCCACCAATGCTTACGCTTA 724
 7626 AGCTGCCATCAACAGCGCTTGAGCAATCTTTGCTGCTGATCCACCAATGCTTACGCTTA 7685
 725 CCAATCCCGGACGCGAGCGGACAGCGGACAGAAAGGTCACTTTGACAGCTGCAATTC 784
 7686 CCAATCCCGGACGCGAGCGGACAGCGGACAGAAAGGTCACTTTGACAGCTGCAATTC 7745
 785 TGGACATCACTACAGAGAGTGTCTCAAGAGATGAAGCGGAGGCTCCACAGTTAAG 844
 7746 TGGACATCACTACAGAGAGTGTCTCAAGAGATGAAGCGGAGGCTCCACAGTTAAG 7805
 845 CTAAGCTTCTATCAATGATGAAGAAAGCTGCAAGCTGACGCCCTCACTTGGGCCAATCTTA 904
 7806 CCAAACTTCTATCTATGATGAAGAAAGCTGCAAGCTGACGCCCTCACTTGGGCCAATCTTA 7865
 905 AATTTGCTATGAGGAGCAAGAGAGCTGCGGAACTTATCCAGCAAGGCAATTAACCATATCC 964
 7866 AATTTGCTATGAGGAGCAAGAGAGCTGCGGAACTTATCCAGCAAGGCAATTAACCATATCC 7925
 965 GCTCCGTGTGGAGAGACTTGTGGAAGACATGAAACACCAATTGACACCAATCATGAG 1024
 7926 GCTCCGTGTGGAGAGACTTGTGGAAGACATGAAACACCAATTGACACCAATCATGAG 7985
 1025 CAAAAAATGAGGTTTCTGCTCCGTCACACGAGAGAGAGCGGCAAGCCAGCTCGCTTA 1084
 7986 CAAAAAATGAGGTTTCTGCTCCGTCACACGAGAGAGAGCGGCAAGCCAGCTCGCTTA 8045
 1085 TCGTGTCCAGACTTGGGGGTCCTGCTGAGGAGAAATGGCCCTTAATGAGTGTGCT 1144
 8046 TCGTGTCCAGACTTGGGGGTCCTGCTGAGGAGAAATGGCCCTTAATGAGTGTGCT 8105
 1145 CCACTCCCTCAGGCTGTGATGAGGCTCTCTGATCGAATTCAGATTCATCTCTGACAGC 1204
 8106 CCACTCCCTCAGGCTGTGATGAGGCTCTCTGATCGAATTCAGATTCATCTCTGACAGC 8165
 1205 GGGTGAAGTCTCTGTGAAGCGCTGGAATCAAGAAAGCCCTTATGAGCTTGTGATAG 1264
 8166 GGGTGAAGTCTCTGTGAAGCGCTGGAATCAAGAAAGCGCTTATGAGCTTGTGATAG 8225
 1265 ACACCGGCTGTTTGACTGACAGTCACTGAGATGACATCCGTGTTGAGAGTCAATAT 1324
 8226 ACACCGGCTGTTTGACTGACAGTCACTGAGATGACATCCGTGTTGAGAGTCAATAT 8285
 1325 ATCAATGTTGATCTTGGCCCCCGAAGCAGACAGGCTTAAGTGTCTACAGAGCGGC 1384
 8286 ACATATGTTGATCTTGGCCCCCGAAGCAGACAGGCTTAAGTGTCTACAGAGCGGC 8345

Qy 1385 TTTAATCGGGGGTCCCTGACTAATTTAAAAAGGCGAATCGCGCTATCGCCGGTGCC 1444
| | | | |
Db 8346 TTTATGTCGGGGGTCCCTGACTAATTTAAAAAGGCGAATCGCGCTATCGCCGGTGCC 8405
| | | | |
Qy 1445 GGGCGAGCGCGTGTCTGAGAGCTAGTGGCGTAAATCCTTCATGTTACTTGAAGGCTT 1504
| | | | |
Db 8406 GGGCGAGCGCGTGTCTGAGAGCTAGTGGCGTAAATCCTTCATGTTACTTGAAGGCTT 8465
| | | | |
Qy 1505 CTGCAAGCTGTGAGAGCTGCAAGAGCTGCAAGAGTGTGCTGTGCGGAGAGCGGC 1564
| | | | |
Db 8466 CTGCGGCTGTGAGAGCTGCAAGAGCTGCAAGAGTGTGCTGTGCGGAGAGCGGC 8525
| | | | |
Qy 1565 TTGCTGTTATCTGTGAGAGCGCGGAAAGCCAGAGAGAGCGCGGAGCTTACAGTCTTCA 1624
| | | | |
Db 8526 TTGCTGTTATCTGTGAAAGCGCGGAAAGCCAGAGAGAGCGCGGAGCTTACAGTCTTCA 8585
| | | | |
Qy 1625 CGGAGGCTATGACTAGTACTTGTGCCCCCGGGGAGACCCGCCCAACCAAGATACGAC 1684
| | | | |
Db 8586 CGGAGGCTATGACTAGTACTTGTGCCCCCGGGGAGACCCGCCCAACCAAGATACGACT 8645
| | | | |
Qy 1685 TGGAGTGTATACATCATGCTCTCCATGTGTGCGTGGAGAGATGATCTGGCAAAA 1744
| | | | |
Db 8646 TGGAGTGTATACATCATGCTCTCTCCATGTGTGCGTGGAGAGATGATCTGGCAAGC 8705
| | | | |
Qy 1745 GGGTATTAAGTCACTGACCCCGTGAACC 1769
| | | | |
Db 8706 GGGTATTAAGTCACTGACCCCGTGAACC 8730
| | | | |
RESULT 15
HPCKR1 9410 bp RNA linear VRL 10-FEB-1999
LOCUS Hepatitis C virus (strain HCV-1b, clone HCV-K1-R1), complete genome
DEFINITION
sequence.
D50480
ACCESSION
D50480.1 GI:1030706
KEYWORDS
polyprotein; core protein; E1; E2; NS2; NS3; NS4A; NS4B; NS5A;
NS5B; envelope protein; non-structural protein;
interferon-sensitive; interferon-resistant; IFN-sensitive;
IFN-resistant; ISDR; interferon sensitivity determining region;
HVR; hypervariable region.
SOURCE
Hepatitis C virus
Hepatitis C virus
ORGANISM
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirinae.
1 (sites)
REFERENCE
AUTHORS
Enomoto, N., Sakuma, I., Asahina, Y., Kurosaki, M., Murakami, T.,
Yamamoto, C., Izumi, N., Marumo, F. and Sato, C.
TITLE
Comparison of full-length sequences of interferon-sensitive and
resistant hepatitis C virus 1b. Sensitivity to interferon is
conferred by amino acid substitutions in the NS5A region
J. Clin. Invest. 96 (1), 224-230 (1995)
JOURNAL
MEDLINE
95340824
PUBMED
7542279
REFERENCE
AUTHORS
Enomoto, N.
JOURNAL
REFERENCE
3 (bases 1 to 9410)
TITLE
Enomoto, N.
Direct Submision
Submitted (08-May-1995) Nobuyuki Enomoto, Tokyo Medical and Dental
University, Second Department of Internal Medicine, 1-5-45 Yushima,
Bunkyo-ku, Tokyo 113, Japan (E-mail: PYN04522@niftyeerve.or.jp,
Tel:03-3813-6111(ex.3224), Fax:03-3818-7177)
FEATURES
source
1..9410
Location/Qualifiers
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/strain="HCV-1b"
/db_xref="taxon:11103"
/clone="HCV-K1-R1"
/note="IFN-resistant HCV from patient 1"
330..9362
CDS

/codon_start=1
/product="polyprotein"
/protein_id="BA09071.1"
/db_xref="GI:1814085"
/translation="MSTRPKPQRTKXNTRNRPDYKPRGGQIVGGVYLLPRGRPL
GVRAIRKTSRQPRGRQPIPKARPERGRMAOPGIPWPIYNGJGAWMLSPRG
SRPMSGPTDPERRRNRLGKVIDLTGCPDLMGIPIVGLAGGAARLAAVLEAD
GVNATGVLNPGCSFSFILALLSLCTIPASVYECNAGLHVINDCSNGLIVEAD
MIMTGGCVPCVRNNSRCAVATLPTLLANANISIPPTTIRRHVDLVLGAAPCSAM
VGLDGSVFLVSCOLTPRSPRYETVPOCNCSIGYHSGHMDMMNPPALAVY
SQLRITQAVYDMAGAKWGLAALAYSNVGNAKYLIMWLEPAGVDGNTYTTGAO
GRVSRILSITPEASQNTIQLINTNGSHIRKTLNCDSDIQTFIALPFAKPNSS
GCPQRLSCRSIDKFAQWGPITYAECHDSQRPCHMYARPGGIVPAQOVCSPYC
FTPSPVVGTTDRFAGPTVMGANEITVLLNNRPPOGMFGCTMNGGTFPTCGG
PCNIGAGANNLTLCPTDCEPKHEATYAKSGSPMLTPRCIVIPYPLMHYPCTFV
SIFVRMYVGVEHRLTAAQNMTRGECDELDRRSLSPILLSTTEMOLLPGFTLL
PALSTGLIHQNIIVDQYLYGVSNVSPRIKREYVILLPLILADARVACIMMLL
VQAQEALENLVNNAASLAGVHGLISFLVFECAMYIKGLVFGAAYALGVPLLL
LLALIPRAVYMDREMAASCGAVPGLALTLSPYKVFALKILMWLYITPAEQ
LQWNIPLNVRGSDAILLTCVAHPELIFDTIKIVALLGLPLMLVQAGIAKVPFR
AOGVIRGCLVRAKVGGOYQOMAEMLKELGTVDHLEPLRDMATGDLAVAYE
PVPSDMEIKITWGGDTAACGDIIMGLPVASRRGRILIGPADSLGGGRLLAPIT
AYSQTRGIRECITITSITGRKDNVBEBOYVSTRATSPLATCNGCWTYTHAGSK
TLAPKGPITQMYTNVDQDLVGNPAPRPAASLTPTCGSDLYLVTHADVIYRRRG
DSRGLSPRIPISTLKSSGGPLCPBGHWVGIIPRAVCTRGVAKAAVDPVESEMTT
MRSPVFTDNSPPPAVPOTFOYAHILAAFTGSKTVPAAVAAQYKVLVNPSTATL
GFGAYMSKAGHGVNDINIRGVRTITGAPITYSTYKGLDGGCGGAYDITMDCES
IDSTIIGITGVLDQAEATAGARVLAAATPPGSAVTPHPHIEIATLSNGEIPFYK
AIPETIKGRGHLIFCHSKKCKDRLAKSLGGLANAVAYRGDLVSVIPASGVVVA
TDLMTGFTGDFDSVIDNTCTVOTVPSLDTFTITFTVPODVAHSORGRTRRG
RGGIYRVTYPERSPGMEFDSVLECEYDAGACAMYELPAETSVARLYLNPGLPFCQ
DHLFPEWGVPLGLTHIDAHPLSQKQAGENPLVIAQATVCARAQAPPPSMDMWC
LTRLKPLTLAGPTPLLYLAGAVONQVTLTHPTTKYIMCMSADLEVTLSYLVGGLA
ALTYACLTGTSVTVIGRIILSGKPAVLPDRALYQEDMEGASHPTYIQQQGLAE
OFKOKALGLQATNQAEEAARPVESKPRALAEAMAKMNMTISGCIAGLISLTPGN
PAKSLAFTASTISPLTGSTLLPILGCVAAQALAPRAVAFVAGAGIAGAVGSI
GLGVLVDTIAGYAGAVAGALVAVKVSGENPSTEDLVNLLPALSPLALVAGVCAA
ILRHVRPGECAVQOMNRLIAFARGHNVSTHVPYSDAAARTQILSLSTIQLR
RLHMDVNDCTPCSGSMLRDVMDMICVLIDPFTMQSKLLPLPGVPFSCRGYR
GVMGDDGIMOTTCPCGCAOITGHNVNGSMRIYGPRTGNTMHTGTPINAYTTGPGTSP
APNYSBALMRVAAEYVETRVGDPHYVTGNTDNYVCPQVPPPEPTVDGRIHR
NAPACGLPRLREVTFOYGLNOYLVGSQPCPEPDDVTVLSMLDPSHITBEAKRL
ARGSPSLASSASSQSLASPLKATCTTHHSPDLDLILANILMQEKGNTTRVSEN
KILILDSFDLPRAEDEREVSIPAEILRKSSKPPALPIVAPRYNPLLSMYODPY
VPPVHGCPLPEPAAPPIPPRRKRTVLTESYTSALAEIATLTPSSGSEADRG
ATAPDQASNDGADGSDVGSYSSWPEEGEPGDDLDLAEASTVSEAGEDEVCCMS
YTWGALITPAAESKLPINALNSILRHNMVYATTSASAOROKVTPDRLQVLD
DHVADLEKMAKSTVRAKLSVEBKRLTPPSAKSRKRGYKADRVNNSKALYD
RSVKDLLEDTEPIDTTIMAKNVEFCVQPEKGRKARLIVFDLIVRCEKALYD
VSTLTPQAVMSSYGFQYSPQORVEPLVNMWKSCKCMGFWYDRCDSTVTESDIV
EESIYOCCLAPPEAROARSLTERLYIGGPLTNKQNGCRRCRAGSVLTSCGNTL
TCYKLSAAGCAALADQCTMLVCGDDLVIVICESAGTOEDDAASLVEPMAEMRYAPAG
DPPEPYVDLITSCSSNVSAHDAASKRYVYLRDPTPLAASLAKMTASTPNSWL
GNITFAPTLVNRMLLMTFPSIILAOQLEKALDQCIYGACISIEPLDIPOTIOLH
GLSAFSLHSYSGEINRVASCLRLGVPLRTMHRARSVAKLLSGGRATLVGRYL
FNMAVVKTKLKLTPIPAASQDLDSNMVFAVYSGGVYVSLSRABRPMFLCLILSVG
GYLLPNR"
330..902
/product="core protein"
903..1478
/product="E1"
1479..2756
/product="E2"
2757..3407
/product="NS2"
3408..5300
/product="NS3"
5301..5462
/product="NS4A"
5463..6245
/product="NS4B"
6246..7586
/product="NS5A"

mac_peptide 7587..9359
 /product="NSSB"
 misc_feature 1479..1559
 /note="hypervariable region (HVR)"
 misc_feature 6934..7073
 /note="interferon sensitivity determining region (ISDR)"
 ORIGIN

Query Match 87.9%; Score 1556.2; DB 14; Length 9410;
 Best Local Similarity 92.5%; Pred. No. 0;
 Matches 1636; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

1 CAATAATGATCTTCCAGAGCTGACCTCATGAGGCGCACTCCGTGGCGGCAATGAGATG 60
 Db CATATGATCTTCCAGAGCTGACCTCATGAGGCGCACTCCGTGGCGGCAATGAGATG 7040
 61 GGGGGGACATTAACCGCGTGAATGAGAGAAAGGTAATTCCTGATCTTTTCGAC 120
 Db GGGGGGACATTAACCGCGTGAATGAGAGAAAGGTAATTCCTGATCTTTTCGAC 7100
 121 CCGCTCCGAGCGGAGAGATGAGCGGGAAGTCCGTCGCGGAGATCTTCGCGGAA 180
 Db CCGCTCCGAGCGGAGAGATGAGCGGGAAGTCCGTCGCGGAGATCTTCGCGGAA 7160
 181 TCCAGAAATTTCCACAGCGATGCCCGCATGGGCAAGCCGGATTACACCTCCGCTG 240
 Db TCCAGAAATTTCCACAGCGATGCCCGCATGGGCAAGCCGGATTACACCTCCGCTG 7161
 241 CTGAGATCTGAGAGGCGCGGAGCTAGTCCCTCCAGATGATGAGATGAGATGAGATG 300
 Db CTGAGATCTGAGAGGCGCGGAGCTAGTCCCTCCAGATGATGAGATGAGATGAGATG 7220
 7221 CTGAGATCTGAGAGGCGCGGAGCTAGTCCCTCCAGATGATGAGATGAGATGAGATG 7280
 301 CCACTAAGACCCCTCTTATACCACTCCAGAGAAAGAGACATGTTCTTTCAGAGAA 360
 Db CCACTAAGACCCCTCTTATACCACTCCAGAGAAAGAGACATGTTCTTTCAGAGAA 7281
 7281 CCGCCAAAGGCGCGCTCAATACCACTCCAGAGAAAGAGACATGTTCTTTCAGAGAA 7340
 361 TCCAGCGATCTTCTGCGCTGAGGAGCTGCGCAAAAGGCTTTGGTAGCTCCGAGCG 420
 Db TCCAGCGATCTTCTGCGCTGAGGAGCTGCGCAAAAGGCTTTGGTAGCTCCGAGCG 7341
 7341 TCCAGCGATCTTCTGCGCTGAGGAGCTGCGCAAAAGGCTTTGGTAGCTCCGAGCG 7400
 421 TCCAGCGATCTTCTGCGCTGAGGAGCTGCGCAAAAGGCTTTGGTAGCTCCGAGCG 480
 Db TCCAGCGATCTTCTGCGCTGAGGAGCTGCGCAAAAGGCTTTGGTAGCTCCGAGCG 7401
 7401 TCCAGCGATCTTCTGCGCTGAGGAGCTGCGCAAAAGGCTTTGGTAGCTCCGAGCG 7460
 481 GCGAGATCTGAGAGGCGCGGAGCTGCGCAAAAGGCTTTGGTAGCTCCGAGCG 540
 Db GCGAGATCTGAGAGGCGCGGAGCTGCGCAAAAGGCTTTGGTAGCTCCGAGCG 7461
 7461 GCGAGATCTGAGAGGCGCGGAGCTGCGCAAAAGGCTTTGGTAGCTCCGAGCG 7520
 541 CCGAGATCTGAGAGGCGCGGAGCTGCGCAAAAGGCTTTGGTAGCTCCGAGCG 600
 Db CCGAGATCTGAGAGGCGCGGAGCTGCGCAAAAGGCTTTGGTAGCTCCGAGCG 7521
 7521 CCGAGATCTGAGAGGCGCGGAGCTGCGCAAAAGGCTTTGGTAGCTCCGAGCG 7580
 601 TCGTGTGATGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
 Db TCGTGTGATGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 7581
 7581 TCGTGTGATGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 7640
 661 AGCAAGCTGCGCATCAACGCGTTGAGCACTTTTGTGCGTCAACCAATGATGCTTAC 720
 Db AGCAAGCTGCGCATCAACGCGTTGAGCACTTTTGTGCGTCAACCAATGATGCTTAC 7641
 7641 AGCAAGCTGCGCATCAACGCGTTGAGCACTTTTGTGCGTCAACCAATGATGCTTAC 7700
 721 GCTACACATCCGCGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 780
 Db GCTACACATCCGCGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 7701
 7701 GCTACACATCCGCGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 7760
 781 ATCTGAGAGATCACTACAGAGAGTGTCAAGAGATGAAAGGCGGCGTCCACAGTT 840
 Db ATCTGAGAGATCACTACAGAGAGTGTCAAGAGATGAAAGGCGGCGTCCACAGTT 7761
 7761 ATCTGAGAGATCACTACAGAGAGTGTCAAGAGATGAAAGGCGGCGTCCACAGTT 7820
 841 AAGGCTAAGCTTTATATAGTAAAGAGAGCTGCAAGTGAAGCGGCCCACTTGGCGCAA 900
 Db AAGGCTAAGCTTTATATAGTAAAGAGAGCTGCAAGTGAAGCGGCCCACTTGGCGCAA 7821
 7821 AAGGCTAAGCTTTATATAGTAAAGAGAGCTGCAAGTGAAGCGGCCCACTTGGCGCAA 7880
 901 TCTAATTTGGTATGGGGAAGAGAGCTCCGGAACCTATCCAGCAAGGCATTAACAC 960

Db 7881 TCCAAATTTGGTATGGGGAAGAGAGCTCCGGAACCTATCCAGCGGCGCGTCAACAC 7940
 Qy 961 ATCCGCTCCGTGTGGAGAGACTTGTGGAAGACACTGAAACCAATTTGACACACATC 1020
 Db 7941 ATCCGCTCCGTGTGGAGAGACTTGTGGAAGACACTGAAACCAATTTGACACACATC 8000
 Qy 1021 ATGGCAAAAATGAGGTTTTCGCGTCAACAGAGAGAGAGGCGGCAAGCCAGCTCCG 1080
 Db 8001 ATGGCAAAAATGAGGTTTTCGCGTCAACAGAGAGAGAGGCGGCAAGCCAGCTCCG 8060
 Qy 1081 CTATGCTGTCCAGACTTGGGAGTCCGTGTGTGCGAGAAAATGAGCCCTTATGACGTG 1140
 Db 8061 CTATGCTGTCCAGACTTGGGAGTCCGTGTGTGCGAGAAAATGAGCCCTTATGACGTG 8120
 Qy 1141 GTCTCAACCTCCCTCAGGCTGTGATGAGGCTCCTGTACGATTCAGATTTCTCTGA 1200
 Db 8121 GTCTCAACCTCCCTCAGGCTGTGATGAGGCTCCTGTACGATTCAGATTTCTCTGA 8180
 Qy 1201 CAGCGGATGAGTCTGCTGTAAGAGCGCTGTAAGTCAAGAGAGCCCTATGGGCTTTGA 1260
 Db 8181 CAGCGGATGAGTCTGCTGTAAGAGCGCTGTAAGTCAAGAGAGCCCTATGGGCTTTGA 8240
 Qy 1261 TATGACACCGGCTGTTTGACTCAACAGTCACTGAGAAATGACATCCGTGTAAGAGATG 1320
 Db 8241 TATGACACCGGCTGTTTGACTCAACAGTCACTGAGAAATGACATCCGTGTAAGAGATG 8300
 Qy 1321 ATTATCAATGTTGTGACTTGGCGCCCGAGCGAGCGAGCCATTAAGTGTCTCACAGAG 1380
 Db 8301 ATTATCAATGTTGTGACTTGGCGCCCGAGCGAGCGAGCCATTAAGTGTCTCACAGAG 8360
 Qy 1381 CCGCTTATATGAGGAGTCCCTGACTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
 Db 8361 CCGCTTATATGAGGAGTCCCTGACTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAG 8420
 Qy 1441 TGGCGGAG 1500
 Db 8421 TGGCGGAG 8480
 Qy 1501 GCGCTGAG 1560
 Db 8481 GCGCTGAG 8540
 Qy 1561 GCGCTGAG 1620
 Db 8541 GCGCTGAG 8600
 Qy 1621 TTCAAG 1680
 Db 8601 TTCAAG 8660
 Qy 1681 GACCTGAG 1740
 Db 8661 GACCTGAG 8720
 Qy 1741 AAAAGGTTATCTACTTCAACCGGTGACC 1769
 Db 8721 AAAAGGTTATCTACTTCAACCGGTGACC 8749

Search completed: February 19, 2005, 07:08:41
 Job time : 7874.71 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2005, 20:55:25 ; Search time 990.878 Seconds
(without alignments)
10574.404 Million cell updates/sec

Title: US-09-664-363-3

Sequence: 1 CAAATGACTTCACGAGCC.....ACTACTCAGCCGATACCG 1770

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980a:*
2: geneseqn1990a:*
3: geneseqn2000a:*
4: geneseqn2001a:*
5: geneseqn2001b:*
6: geneseqn2002a:*
7: geneseqn2002b:*
8: geneseqn2003a:*
9: geneseqn2003b:*
10: geneseqn2003c:*
11: geneseqn2003d:*
12: geneseqn2004a:*
13: geneseqn2004b:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1770	100.0	1770	2	AAQ12236 Clone JG2
2	1762	99.5	3750	2	AAQ12241 Encodes P
3	1553.8	87.8	8638	6	ABK88596 Hepatitis
4	1553.8	87.8	8638	6	ABK88578 Hepatitis
5	1553.4	87.8	8642	6	ABK88574 Hepatitis
6	1552.2	87.7	7141	6	AAQ25333 Hepatitis
7	1552.2	87.7	7789	6	AAQ25330 Hepatitis
8	1552.2	87.7	7987	6	AAQ25329 Hepatitis
9	1552.2	87.7	7987	6	AAQ25321 Hepatitis
10	1552.2	87.7	7987	6	AAQ25324 Hepatitis
11	1552.2	87.7	7989	3	AAA8968 Hepatitis
12	1552.2	87.7	7989	6	AAQ25322 Hepatitis
13	1552.2	87.7	7989	6	AAQ25326 Hepatitis
14	1552.2	87.7	7989	6	AAQ25325 Hepatitis
15	1552.2	87.7	7989	10	ADD93722 Hepatitis
16	1552.2	87.7	7989	10	ADD93734 Hepatitis
17	1552.2	87.7	7991	12	ADJ57845 HCV repli
18	1552.2	87.7	7992	6	AAQ17279 Hepatitis
19	1552.2	87.7	7992	6	AAQ17276 Hepatitis
20	1552.2	87.7	7992	6	AAQ17280 Hepatitis

21	1552.2	87.7	7992	6	AAQ17277 Hepatitis
22	1552.2	87.7	7992	6	AAQ17281 Hepatitis
23	1552.2	87.7	7995	6	AAQ17278 Hepatitis
24	1552.2	87.7	8001	3	AAA89667 Hepatitis
25	1552.2	87.7	8637	3	AAA89686 Hepatitis
26	1552.2	87.7	8638	6	ABK88595 Hepatitis
27	1552.2	87.7	8638	6	ABK88577 Hepatitis
28	1552.2	87.7	8639	6	ABK88573 Hepatitis
29	1552.2	87.7	8643	6	ABK88575 Hepatitis
30	1552.2	87.7	8649	3	AAA89699 Hepatitis
31	1552.2	87.7	8732	10	ADQ48121 RNA HCV r
32	1552.2	87.7	9605	6	ABK91431 Hepatitis
33	1552.2	87.7	9605	6	ABK91424 Hepatitis
34	1552.2	87.7	9605	6	ABK91429 Hepatitis
35	1552.2	87.7	9605	6	ABK91432 Hepatitis
36	1552.2	87.7	9605	6	ABK91411 Hepatitis
37	1552.2	87.7	9605	6	ABK91430 Hepatitis
38	1552.2	87.7	9605	6	ABK91428 Hepatitis
39	1552.2	87.7	9605	6	ABK91425 Hepatitis
40	1552.2	87.7	9605	6	ABK91426 Hepatitis
41	1552.2	87.7	9605	6	ABK91433 Hepatitis
42	1552.2	87.7	9605	6	AAQ25332 Hepatitis
43	1552.2	87.7	9608	6	ABK91427 Hepatitis
44	1552.2	87.7	10690	6	ABK91448 Hepatitis
45	1552.2	87.7	10690	6	ABK91435 Hepatitis

ALIGNMENTS

RESULT 1
AAQ12236
ID AAQ12236 standard; cDNA to rRNA; 1770 BP.
XX
AC AAQ12236;
XX
AC
XX
DT 25-MAR-2003 (revised)
DT 06-FEB-1991 (first entry)
DE Clone JG2 encoding portion of PT-NANB polypeptide.
XX
XX post-transfusional non-A, non-B hepatitis; virus; vaccine; ss.
XX
OS Non-A.
OS non-B hepatitis virus.
OS
PN GB2239245-A.
XX
XX 26-JUN-1991.
PD
XX
XX 17-DEC-1990; 90GB-00027250.
PF
XX
PR 18-DEC-1989; 89GB-00028562.
PR 27-FEB-1990; 90GB-00004414.
PR 03-MAR-1990; 90GB-00004814.
XX
XX (WELLS) WELLCOME FOUND LTD.
PA (HIGH/) HIGHFIELD P E.
XX
PI Highfield PE, Rodgers BC, Tedder RS, Barbara JAV;
XX
DR WPI; 1991-187584/26.
DR P-PSDB; AAR12594.
XX
PT Post-transfusional non-A non-B hepatitis poly:peptide(s) - and also DNA
PT and antibodies used in diagnostic assays and in vaccines.
XX
PS Claim 10; Page 48-52; 108pp; English.
XX
XX This sequence encodes a portion of the PT-NANB polypeptide. It was
CC isolated from serum of humans infected by the virus. Genomic RNA from
CC viruses pelleted from the infected serum was used to prepare a cDNA
CC library in lambda gtl1 which was screened with antisera and

CC oligonucleotide probes. One of the positive plaques to be obtained was
 CC JG2. DNA was extracted from denatured phage stocks, amplified by PCR and
 CC sequenced. See also AA012237-Q12242. (Updated on 25-MAR-2003 to correct
 CC PA field.)

XX Sequence 1770 BP; 402 A; 531 C; 496 G; 341 T; 0 U; 0 Other;

Query Match 100.0%; Score 1770; DB 2; Length 1770;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CAAATGACTTCCGAGACGCTGACTCATGAGGCCAACCTCTGTGGCGGATGAGT 60
DB 1 CAAATGACTTCCGAGACGCTGACTCATGAGGCCAACCTCTGTGGCGGATGAGT 60
QY 61 GCGGGGACATTACCGCGGTGAGTCAAGAAACAAGTAGTAATCTGTGACTTTCCAC 120
DB 61 GCGGGGACATTACCGCGGTGAGTCAAGAAACAAGTAGTAATCTGTGACTTTCCAC 120
QY 121 CCGCTCCGAGCGGAGAGATGAGCGGGAAGTCCCTCCGCGGAGATCCTGCGGAAA 180
DB 121 CCGCTCCGAGCGGAGAGATGAGCGGGAAGTCCCTCCGCGGAGATCCTGCGGAAA 180
QY 181 TCCAGAAATTTCCACACAGCATGCTCCGATGCGGACGCGGATTAACAACCTTCGCTG 240
DB 181 TCCAGAAATTTCCACACAGCATGCTCCGATGCGGACGCGGATTAACAACCTTCGCTG 240
QY 241 CTGAGTCTCTGAGAGGCGCGGACCTACGTCCTCCAGTGTGATCATGGGTGCCCATCTGCA 300
DB 241 CTGAGTCTCTGAGAGGCGCGGACCTACGTCCTCCAGTGTGATCATGGGTGCCCATCTGCA 300
QY 301 CCTACTAAGACCCCTCTATATACACCTCCACGGAAGAAAGAGACAGTTGTTGACAGAA 360
DB 301 CCTACTAAGACCCCTCTATATACACCTCCACGGAAGAAAGAGACAGTTGTTGACAGAA 360
QY 361 TCCACCGTGTCTTTCGCTTGGCGGAGCTTGGCCACAAAGGCTTTTGGTAGTCTCCGACCG 420
DB 361 TCCACCGTGTCTTTCGCTTGGCGGAGCTTGGCCACAAAGGCTTTTGGTAGTCTCCGACCG 420
QY 421 TCGGCGCTGACAGCGGCGACCGGCAACCGCCCTCCGACCAATCCCTCCGACGACGCGGCA 480
DB 421 TCGGCGCTGACAGCGGCGACCGGCAACCGCCCTCCGACCAATCCCTCCGACGACGCGGCA 480
QY 481 GCAGGATCTGACGTTGAGTGTGATTCCTCCATGCCCTTGAAGGGGAGCGGGGAC 540
DB 481 GCAGGATCTGACGTTGAGTGTGATTCCTCCATGCCCTTGAAGGGGAGCGGGGAC 540
QY 541 CCCGATCTCAGCGAGCGGTCTTGGTCTACCGTAGAGAGGCGGCTGAGGACGTCGTC 600
DB 541 CCCGATCTCAGCGAGCGGTCTTGGTCTACCGTAGAGAGGCGGCTGAGGACGTCGTC 600
QY 601 TSCCTCGATGCTCCCAACATGAGCAGGCGCTGATCAGGCAATGGCGTCCGAGGAA 660
DB 601 TSCCTCGATGCTCCCAACATGAGCAGGCGCTGATCAGGCAATGGCGTCCGAGGAA 660
QY 661 AGCAAGCTGCCCATCAACGCGTTGAGCAACTTTGTGCTGCGTCAACACAATGCTCTAC 720
DB 661 AGCAAGCTGCCCATCAACGCGTTGAGCAACTTTGTGCTGCGTCAACACAATGCTCTAC 720
QY 721 GCTACCAATCCCGAGCGGCAAGCGGCGGCAAGAAAGTACCTTTGACAGACTGCA 780
DB 721 GCTACCAATCCCGAGCGGCAAGCGGCGGCAAGAAAGTACCTTTGACAGACTGCA 780
QY 781 ATCTGAGAGATCACTACAGGACGCTCAAGAGATGAAGGCGACGCTCCACAGTT 840
DB 781 ATCTGAGAGATCACTACAGGACGCTCAAGAGATGAAGGCGACGCTCCACAGTT 840
QY 841 AAGGCTAAGCTTCTATCAGTAGAGAGAGCTGCAAGCTCCCACTTTCCGCGCAA 900
DB 841 AAGGCTAAGCTTCTATCAGTAGAGAGAGCTGCAAGCTCCCACTTTCCGCGCAA 900
QY 901 TCTAATTTGGTATGGGCAAGAGAGCTCCGGAACCTATCCAGCAAGGCCATTACCAAC 960
DB 901 TCTAATTTGGTATGGGCAAGAGAGCTCCGGAACCTATCCAGCAAGGCCATTACCAAC 960

```

```

DB 901 TCTAATTTGGTATGGGCAAGAGAGCTCCGGAACCTATCCAGCAAGGCCATTACCAAC 960
QY 961 ATCCGCTCCGCTGGGAGAGACTTTGTTGGAAGACATTGAAAACCAATTGACACCCATC 1020
DB 961 ATCCGCTCCGCTGGGAGAGACTTTGTTGGAAGACATTGAAAACCAATTGACACCCATC 1020
QY 1021 ATGGCAAAAATGAGGTTTTCGCTCCACAGAGAGAGAGCGCGCAAGCACTGAC 1080
DB 1021 ATGGCAAAAATGAGGTTTTCGCTCCACAGAGAGAGAGCGCGCAAGCACTGAC 1080
QY 1081 CTATGCTGTTCCAGACTTGGGAGGTCGTGTGTGCGAAGAAATGAGCCCTTATGACGTG 1140
DB 1081 CTATGCTGTTCCAGACTTGGGAGGTCGTGTGTGCGAAGAAATGAGCCCTTATGACGTG 1140
QY 1141 GTCTCACCCTCCCTCAAGCTGTGATGAGGCTCCCTGTACAGATTCCATTTCTCTGGA 1200
DB 1141 GTCTCACCCTCCCTCAAGCTGTGATGAGGCTCCCTGTACAGATTCCATTTCTCTGGA 1200
QY 1201 CAGCGGCTCGAGTTCTGTGTAACGCTGGAATCAAGAAAGACCCCTATGGGCTTTGCA 1260
DB 1201 CAGCGGCTCGAGTTCTGTGTAACGCTGGAATCAAGAAAGACCCCTATGGGCTTTGCA 1260
QY 1261 TATGACACCCGCTGTTTGACTCAACAGTCACTGAGATGACATCCGTAGAGAGTCA 1320
DB 1261 TATGACACCCGCTGTTTGACTCAACAGTCACTGAGATGACATCCGTAGAGAGTCA 1320
QY 1321 ATTATCAATGTGATCTTGGGCGCGGACGACGAGGCAATPAGGTCCCTCAACAG 1380
DB 1321 ATTATCAATGTGATCTTGGGCGCGGACGACGAGGCAATPAGGTCCCTCAACAG 1380
QY 1381 CGGCTTTATATCGGAGGTCCTCTGACTTAATCAAAAGGAGCAAGTGGGCTATCGCCG 1440
DB 1381 CGGCTTTATATCGGAGGTCCTCTGACTTAATCAAAAGGAGCAAGTGGGCTATCGCCG 1440
QY 1441 TCCCGCGAGCGGCTGTCTGACGACTGCGGTAAATCCCTACATGTTACTTGAAG 1500
DB 1441 TCCCGCGAGCGGCTGTCTGACGACTGCGGTAAATCCCTACATGTTACTTGAAG 1500
QY 1501 GCTCTGAGGCTGTGAGCTGCAAGAGCTCGAGACTGCAAGATGCTGTGTGGGAGAC 1560
DB 1501 GCTCTGAGGCTGTGAGCTGCAAGAGCTCGAGACTGCAAGATGCTGTGTGGGAGAC 1560
QY 1561 GGCCTGTGCTTATCTGTGAGAGCGCGGAAACCAAGAGAGAGCGGCGACTGACGATC 1620
DB 1561 GGCCTGTGCTTATCTGTGAGAGCGCGGAAACCAAGAGAGAGCGGCGACTGACGATC 1620
QY 1621 TTCAAGAGGCTATGACTGAGTACTTGCCTCCGCGGAGACCCGCCAACAAGATAC 1680
DB 1621 TTCAAGAGGCTATGACTGAGTACTTGCCTCCGCGGAGACCCGCCAACAAGATAC 1680
QY 1681 GACCTGAGTGTGATTAACATCATGCTCCCAATGTGTGCGTGGGCAAGATGCTGCGC 1740
DB 1681 GACCTGAGTGTGATTAACATCATGCTCCCAATGTGTGCGTGGGCAAGATGCTGCGC 1740
QY 1741 AAAAGGATATCTACTCTCAACCCGTGACCG 1770
DB 1741 AAAAGGATATCTACTCTCAACCCGTGACCG 1770

```

RESULT 2
 AA012241
 ID AA012241 standard; cDNA, 3750 BP.

XX AA012241;
 XX AC
 XX 25-MAR-2003 (revised)
 DT 17-SEP-1991 (first entry)
 XX DE
 XX Encodes portion of PT-NANBH viral non-structural protein.
 KM post-transfusional non-A, non-B hepatitis; virus; vaccine; ss.
 OS Non-A.

OS non-B hepatitis virus.
XX GB2239245-A.
XX PD 26-JUN-1991.
XX PF 17-DEC-1990; 90GB-00027250.
XX PR 18-DEC-1989; 89GB-00028562.
XX PR 27-FEB-1990; 90GB-0000414.
XX PR 03-MAR-1990; 90GB-00004814.
XX PA (WELL) WELLCOME FOUND LTD.
XX PA (HIGH/) HIGHFIELD P E.
XX PI Highfield PE, Rodgers BC, Tedder RS, Barbara JAJ;
XX DR WPI; 1991-187584/26.
XX DR P-PSDB; AAR12599.
XX PT Post-transfusal non-A non-B hepatitis poly:peptide(s) - and also DNA
XX PT and antibodies used in diagnostic assays and in vaccines.
XX PS Claim 10; Page 88-97; 108pp; English.
XX CC This sequence probably encodes viral non-structural proteins of the PT-
CC NANBH viral genome which are antigenic. It was isolated from serum of
CC humans infected by the virus. See also AAQ12236-40 and AAQ12242. (Updated
CC on 25-MAR-2003 to correct PA field.)
XX SO Sequence 3750 BP; 794 A; 1140 C; 1072 G; 744 T; 0 U; 0 Other;
Query Match 99.5%; Score 1762; DB 2; Length 3750;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1765; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 CAAATGACTTCCAGACGGTGAAGCTTCATGAGGCACTCCGTGGCGGAGTGAAGT 60
DB 1981 CAAATGACTTCCAGACGGTGAAGCTTCATGAGGCACTCCGTGGCGGAGTGAAGT 2040
QY 61 GGGGGGACATTACCCCGGTGAGTCAAGAAAGAGTAACTCTGACTCTTTTCGAC 120
DB 2041 GGGGGGACATTACCCCGGTGAGTCAAGAAAGAGTAACTCTGACTCTTTTCGAC 2100
QY 121 CCGCTCCGAGCGGAGAGTGAAGCGGAAAGTGTCCGTCCCGGCGAGATCTCGGAAA 180
DB 2101 CCGCTCCGAGCGGAGAGTGAAGCGGAAAGTGTCCGTCCCGGCGAGATCTCGGAAA 2160
QY 181 TCCAGAAATTCCACAGGATGCCGCAATGGGCAAGCCGGAATTACACCTCCGCTG 240
DB 2161 TCCAGAAATTCCACAGGATGCCGCAATGGGCAAGCCGGAATTACACCTCCGCTG 2220
QY 241 CTGAGATCTGGAAGGCGCGGACTAGTCCCTCCAGTGTATCATGGGTCCCACTGCCA 300
DB 2221 CTGAGATCTGGAAGGCGCGGACTAGTCCCTCCAGTGTATCATGGGTCCCACTGCCA 2280
QY 301 CCTACTAAGACCCCTCTCTATACCACTTCAACGAGAAAGAGACAGTTGTTCTGACAAA 360
DB 2281 CCTACTAAGACCCCTCTCTATACCACTTCAACGAGAAAGAGACAGTTGTTCTGACAAA 2340
QY 361 TCCAGCGTGTCTCTGCGCTGGGAGGCTTGCACAAAGGCTTTGGTACTCCGACCG 420
DB 2341 TCCAGCGTGTCTCTGCGCTGGGAGGCTTGCACAAAGGCTTTGGTACTCCGACCG 2400
QY 421 TCGGCGGTGACAGCGGCAAGGCAAGCCCTCTGACCAATCTTCCGACGAGCGGGA 480
DB 2401 TCGGCGGTGACAGCGGCAAGGCAAGCCCTCTGACCAATCTTCCGACGAGCGGGA 2460
QY 481 GCGAGATCTGAGCTGTGATTTCTCTCATAGCCCCCTTTGAGGGGAGCGGGGAGAC 540
DB 2461 GCGAGATCTGAGCTGTGATTTCTCTCATAGCCCCCTTTGAGGGGAGCGGGGAGAC 2520
QY 541 CCGAGATCTGAGCGAGGCTTGTGTCTACGATGATGAGAGCGCGGTGAGAGCTGTCTC 600

DB 2521 CCGAGATCTGAGCGAGGCTTGTGTCTACGATGATGAGAGCGCGGTGAGAGCTGTCTC 2580
QY 601 TGCTGCTGATGTCTTACACATGACAGGCGCTCTGATACACCGCATGCGCTGGAGGAA 660
DB 2581 TGCTGCTGATGTCTTACACATGACAGGCGCTCTGATACACCGCATGCGCTGGAGGAA 2640
QY 661 AGCAAGCTGCCATTCACAGCGGTTGAGCAACTCTTGTGCTGCTGACACCAACATGCTAC 720
DB 2641 AGCAAGCTGCCATTCACAGCGGTTGAGCAACTCTTGTGCTGCTGACACCAACATGCTAC 2700
QY 721 GCTACCAATCCCGCAGCGCAAGCCAGCGGAGAGAGGTCACCTTTGACAGACTGCAA 780
DB 2701 GCTACCAATCCCGCAGCGGCAAGCCAGCGGAGAGAGGTCACCTTTGACAGACTGCAA 2760
QY 781 ATCTGAGAGATCACTACAGAGAGCTGTCTACAGAGATGAGAGCGGCTCACAGTT 840
DB 2761 ATCTGAGAGATCACTACAGAGAGCTGTCTACAGAGATGAGAGCGGCTCACAGTT 2820
QY 841 AAGGCTAAGCTTCTACAGTGAAGAGCGCTGCAAGCTGACCGCCCACTTCGCGCAAA 900
DB 2821 AAGGCTAAGCTTCTACAGTGAAGAGCGCTGCAAGCTGACCGCCCACTTCGCGCAAA 2880
QY 901 TCTAATTTGGCTATGAGGCAAAAGAGCTCCGGAACCTATCCAGCAAGGCAATTAAACAC 960
DB 2881 TCTAATTTGGCTATGAGGCAAAAGAGCTCCGGAACCTATCCAGCAAGGCAATTAAACAC 2940
QY 961 ATCCGCTCGTGTGAGAGAGCTTTGTTGGAAGACCTGAAACCAATTGACACCACTATC 1020
DB 2941 ATCCGCTCGTGTGAGAGAGCTTTGTTGGAAGACCTGAAACCAATTGACACCACTATC 3000
QY 1021 ATGGCAAAATGAGGTTTTTTCGCGTCAACAGAGAGAGGCGGCAAGCCAGCTCCG 1080
DB 3001 ATGGCAAAATGAGGTTTTTTCGCGTCAACAGAGAGAGGCGGCAAGCCAGCTCCG 3060
QY 1081 CTATGCTGTCCACACTTGGGGGTCGCTGTGTGCAAGAAATGAGCCCTTATGACGTG 1140
DB 3061 CTATGCTGTCCACACTTGGGGGTCGCTGTGTGCAAGAAATGAGCCCTTATGACGTG 3120
QY 1141 GTCTCAACCTCTCTCAGGCTGTGATGAGCTCTCTCTGATTCAGATTCAGATTCCTGGA 1200
DB 3121 GTCTCAACCTCTCTCAGGCTGTGATGAGCTCTCTCTGATTCAGATTCAGATTCCTGGA 3180
QY 1201 CAGCGGATGAGGTTCTGTGTAAGCGCTGGAATCAAAAGCCCTTATGAGCTTTGGA 1260
DB 3181 CAGCGGATGAGGTTCTGTGTAAGCGCTGGAATCAAAAGCCCTTATGAGCTTTGGA 3240
QY 1261 TATGACCCGCTGTTTGAATCAAGTCACTGAGATGACATCCGTGTAGAGAGTCA 1320
DB 3241 TATGACCCGCTGTTTGAATCAAGTCACTGAGATGACATCCGTGTAGAGAGTCA 3300
QY 1321 ATTATCAATGTTGTGACTTGGCCCCCGAAGCAGACAGGCAATGAGTCCGCTCACAGAG 1380
DB 3301 ATTATCAATGTTGTGACTTGGCCCCCGAAGCAGACAGGCAATGAGTCCGCTCACAGAG 3360
QY 1381 CCGCTTATATCGGGGCTCCCTGACTAATTCAAAAGGCAAGCTGCGGCTATCGCCGG 1440
DB 3361 CCGCTTATATCGGGGCTCCCTGACTAATTCAAAAGGCAAGCTGCGGCTATCGCCGG 3420
QY 1441 TGGCGGAGAGGCGTGTGACAGTACTAGCTGAGTATACCTTCATGATTAATTGAG 1500
DB 3421 TGGCGGAGAGGCGTGTGACAGTACTAGCTGAGTATACCTTCATGATTAATTGAG 3480
QY 1501 GCTCTGACAGCTGTGAGCTGCAAGGCTCAGAGCTGACAGATGCTGTGTGCGAGAC 1560
DB 3481 GCTCTGACAGCTGTGAGCTGCAAGGCTCAGAGCTGACAGATGCTGTGTGCGAGAC 3540
QY 1561 GGCCTGTGTATCTGTGAGAGCGGGGAAACCCAGAGAGAGCGCGGAGCTTACAGAGTC 1620
DB 3541 GGCCTGTGTATCTGTGAGAGCGGGGAAACCCAGAGAGAGCGCGGAGCTTACAGAGTC 3600
QY 1621 TTCAGGAGGCTATGACTAGTACTGTGCCCCCGCGGAGACCGGCCCCAAGCAATAC 1680

Db	3601	TTCAAGGAGGCGTATACCTAGTACTCTGCCCCCCCCGGGAGACCGCCCCCAACGAGAAATAC	3668
Oy	1681	GACCTGGAGGTTGATATACATCATGCTCTCCCAATGTGTGGTGGGCAAGATGCATCTGGC	1740
Db	3661	GACCTGGAGGTTGATATACATCATGCTCTCCCAATGTGTGGTGGGCAAGATGCATCTGGC	3720
Oy	1741	AAAAAGGTATCTACTACCTCAGCCCGTGACCCG	1770
Db	3721	AAAAAGGTATCTACTACCTCAGCCCGTGACCCG	3750
RESULT 3			
ID	ABK88596	ABK88596	
XX	ABK88596	standard; DNA; 8638 BP.	
AC	ABK88596;		
XX			
DT	21-OCT-2002	(first entry)	
XX			
DE	Hepatitis C virus R3-rep-5'A replicon.		
XX			
KW	Self-replicating; hepatitis C virus; HCV; HCV replication inhibitor;		
KM	cell culture replication; ds; mutant.		
XX			
OS	Hepatitis C virus.		
XX			
PH	Key	Location/Qualifiers	
FT	mutation	replace(1,G)	
FT		/*tag= a	
PT	CDS	1802..8407	
PT		/*tag= b	
FT		/product= "HCV NS2-5B"	
FT	mutation	/note= "Viral enzymes"	
FT		replace(2509,T)	
FT		/*tag= c	
FT	mutation	replace(2778,G)	
FT		/*tag= d	
FT	mutation	replace(2840,A)	
FT		/*tag= e	
FT	mutation	replace(3574,T)	
FT		/*tag= f	
FT	mutation	replace(4052,A)	
FT		/*tag= g	
FT	mutation	replace(4387,C)	
FT		/*tag= h	
FT	mutation	replace(4446,A)	
FT		/*tag= i	
FT	mutation	replace(4507,C)	
FT		/*tag= j	
FT	mutation	replace(4855,T)	
FT		/*tag= k	
FT	mutation	replace(5351,A)	
FT		/*tag= l	
FT	mutation	replace(5498,G)	
FT		/*tag= m	
FT	mutation	replace(5659,G)	
FT		/*tag= n	
FT	mutation	replace(5838,T)	
FT		/*tag= o	
FT	mutation	replace(5871,C)	
FT		/*tag= p	
FT	mutation	replace(6115,A)	
FT		/*tag= q	
XX			
PN	WO200252015-A2.		
XX			
PD	04-JUL-2002.		
XX			
PF	20-DEC-2001; 2001WO-CA001843.		
XX			
PR	22-DEC-2000; 2000US-0257857P.		
XX			
XX	(BOEH) BOEHRINGER INGELHEIM CANADA LTD.		

XX	Kukkolj G,	Pause A;
PI		
DR	WPI; 2002-575382/61.	
XX		
XX	New self-replicating RNA molecules from Hepatitis C virus (HCV), which	
PT	possess enhanced transduction or replication efficiency, useful for	
PT	evaluating potential inhibitors of HCV replication.	
PS	Claim 11; Page 130-140; 140pp; English.	
XX		
CC	The invention describes a self-replicating hepatitis C virus (HCV)	
CC	polynucleotide molecule comprising a 5'-non translated region (NTR),	
CC	where guanine at position 1 is substituted for adenine, a HCV polypeptide	
CC	region coding for a HCV polypeptide, and a 3'-NTR region. The self-	
CC	replicating Hepatitis C virus (HCV) RNA molecule is useful for evaluating	
CC	potential inhibitors of HCV replication. The HCV RNA molecule is also	
CC	useful for efficiently establishing cell culture replication. The self-	
CC	replicating polynucleotide molecule contains a 5'-NTR, where G at	
CC	position 1 is substituted for A, and therefore provides an alternative to	
CC	existing systems comprising a self-replicating HCV RNA molecule that, in	
CC	conjunction with mutations in the HCV non-structural region, such as the	
CC	G(2042)C/R mutations, transduces and/or replicates with greater	
CC	efficiency. This sequence represents hepatitis C virus replicon R3-rep-	
CC	5'G, a self-replicating HCV polynucleotide molecule created from the	
CC	APK612 (see ABK8573)	
SQ	Sequence 8638 BP; 1767 A; 2558 C; 2440 G; 1873 T; 0 U; 0 Other;	
Query Match	87.8%; Score 1553.8; DB 6; Length 8638;	
Best Local Similarity	92.5%; Pred. No. 0;	
Matches 1633; Conservative	0; Mismatches 132; Indels 0; Gaps 0	
OY	5	ATGACTTCCAGAGCGCTACCTCATCGAGGCCAACCTCTGTGGCGCATAGATGGCG 64
Db	6030	ATGACTCCCCCGAGCGCTGACTCACTCAGCGGCACCTCTGTGGCGCATAGATGGCG 6088
OY	65	GGAACATTACC CGGTGGAGTCAGAGAACAAGTAGTAATCCTTGACCTTTGACCCGC 124
Db	6090	GGAACATACC CGGTGGAGTCAGAGAATAAGTAGTAATTTTGGACTCTTTTCGAGCCG 6144
OY	125	TCCGAGCGAGAGATGAGCGGAGAGTGTCCGTCCGCGGAGATCTCTGCGAATCCA 184
Db	6150	TCCAAAGCGAGAGATGAGAGGAGATGATCCGTTCCGCGGAGATCTCTGCGAGTCCA 6208
OY	185	AGAAATTCACACAGCATGCCCGCATGGGCAACCGCGGATTAACAACCTTCGCTGCG 244
Db	6210	GGAATTCCTCCAGACGATGCCCATATGGGCAACCGCGGATTAACAACCTTCGCTTGA 6268
OY	245	AGCTCTGGAAGGCCCGGACTACGTCCTCCAGTGGTACATGAGGCGCCACATGCCACTA 304
Db	6270	AGTCTTGGAAGAACCGGACTACGTCCTCCAGTGGTACATGAGGCTCATTTGCCCTTG 6328
OY	305	CTAAGACCCCTCTATACCACTCCACGACGAAAAGAGACAGTTGTTGACAGATCCA 364
Db	6330	CAAAGGCCCTCCGATACCACTCCACGAGGAAGAGAGACGTTGTCTGTACAGATCTA 6388
OY	365	CCGTGTCTTTCGCTTCGCGGAGCTTGCACAAAGGCTTTTGTGTAAGCTTCGACCTGTGG 424
Db	6390	CCGTGTCTTTCGCTTTGGCGAGCTTCGCAAAAGACCTTGGCGAGCTCCGAAATGTGCG 6448
OY	425	CCGTGACAGGGGACGCGCAACCGCCCTCTGACCAATCTCCGACGACGGCGGAGAG 484
Db	6450	CCGTGACAGGGGACGCGCAACCGGCTCTTCTGACCAACCTTCGACGACGCGACGCG 6508
OY	485	GATCTGACGTTGAGTCGTATTCCTTCATGCCCCCTTTGAGGGGAGCGGAGAACCCG 544
Db	6510	GATCCGAGTTGAGTCGTATTCCTTCATGCCCCCTTTGAGGGGAGCGGAGAAATCCG 6568
OY	545	ATCTCAGGAGCGGCTTGTGTCTACCGTGAATGAGAGAGCGGATGAGACGTCGTCTGT 604
Db	6570	ATCTCAGGAGCGGCTTGTGTCTACCGTAACGAGAGAGCGTAAGTGAACGTCGTCTGT 6628

PT New self-replicating RNA molecules from Hepatitis C virus (HCV), which
PT possess enhanced transduction or replication efficiency, useful for
PT evaluating potential inhibitors of HCV replication.

XX Claim 11: Page 106-116; 140pp; English.

XX The invention describes a self-replicating hepatitis C virus (HCV)
CC polynucleotide molecule comprising a 5'-non translated region (NTR),
CC where guanine at position 1 is substituted for adenine, a HCV polyprotein
CC region coding for a HCV polyprotein, and a 3'-NTR region. The self-
CC replicating Hepatitis C virus (HCV) RNA molecule is useful for evaluating
CC potential inhibitors of HCV replication. The HCV RNA molecule is also
CC useful for efficiently establishing cell culture replication. The self-
CC replicating polynucleotide molecule contains a 5'-NTR, where G at
CC position 1 is substituted for A, and therefore provides an alternative to
CC existing systems comprising a self-replicating HCV RNA molecule that, in
CC conjunction with mutations in the HCV non-structural region, such as the
CC G(2042)C/R mutations, transduces and/or replicates with greater
CC efficiency. This sequence represents hepatitis C virus replicon R3-rep, a
CC self-replicating HCV polynucleotide molecule created from the APOK12 (see
CC ABK8573)

XX Sequence 8638 BP, 1766 A, 2558 C, 2441 G, 1873 T, 0 U, 0 Other;

Query Match 87.8%; Score 1553.8; DB 6; Length 8638;

Best Local Similarity 92.5%; Pred. No. 0;

Matches 1633; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

```
QY 5 ATGACTTCCGAGACGCTGATCTGAGGCCAACCTCTGTGGCGGATGAGATGGGCG 64
DB 6030 ATGACTCCCCGAGACGCTGATCTGAGGCCAACCTCTGTGGCGGAGATGGGCG 6089
QY 65 GGGACATTACCCGCGTGAAGTCAAGAGAGAGTAAATCCGAGCTCTTTCGACCCGC 124
DB 6030 ATGACTCCCCGAGACGCTGATCTGAGGCCAACCTCTGTGGCGGAGATGGGCG 6089
QY 6090 GGAATCTACCCGCGTGAAGTCAAGAGAGAGTAAATCCGAGCTCTTTCGACCCGC 6149
DB 125 TCCGACCGAGAGAGATGACGCGGAGATGTCCTCCGCGGAGATCTCGGAAATCCA 184
DB 6150 TCCAGCGGAGAGATGAGAGGAGATGATCCGTTCCGCGGAGATCTCGGAGATCCA 6209
QY 185 AGAATTTCCACACGAGATGCCGATGGGACACGCCGATTAACAACCTCCGCTGCTG 244
DB 6210 GGAATTTCCCTCGAGGATGCCATATGGGACACGCCGATTAACAACCTCCGCTG 6269
QY 245 AGTCCGGAAGGCCCGGACCTAGTCCCTCCAGTGTATAGGAGGCCACTGCCACCTA 304
DB 6270 AGTCTGGAAGGCCCGGACCTAGTCCCTCCAGTGTATAGGAGGCCACTGCCCTG 6329
QY 305 CTAAAGACCCCTCTATACCACTCCACGAGAGAAAGAGACAGTTGTTCTGACAGATCCA 364
DB 6330 CCAAGGCCCTCTCGATACCACTCCACGAGAGAAAGAGACAGTTGTTCTGACAGATCCA 6389
QY 365 CCGTGTCTTCTGCTGCGTGGAGCTTGGCAAAAGCTTTTGGTAGCTCCGACCGTGG 424
DB 6390 CCGTGTCTTCTGCTGCGTGGAGCTTGGCAAAAGCTTTTGGTAGCTCCGATCGTCCG 6449
QY 425 CCGTGACAGCGGACGAGCAACCGGCCCTCTGACCAATCTCCACGAGCGGAGACGAG 484
DB 6450 CCGTGACAGCGGACGAGCAACCGGCCCTCTCTGACCAAGCCCTCCGACGAGCGGAGCGG 6509
QY 485 GATCTGACGTTGATGCTATCTCTCCATGCCCCCTTGAAGGAGGAGCGGGGAGATCCCG 544
DB 6510 GATCTGACGTTGATGCTATCTCTCCATGCCCCCTTGAAGGAGGAGCGGGGAGATCCCG 6569
QY 545 ATCTCAGCGACGGGTTTGGTCTACCTGTAGTGAAGAGCGCGTGAAGACGTCGTCTGCT 604
DB 6570 ATCTCAGCGACGGGTTTGGTCTACCTGTAAAGCAGAGAGCTAGTGAAGACGTCGTCTGCT 6629
QY 605 GGTGATGTCCTACACATGAGAGCGGCTGTATCAAGCATGCGGTGGAGGAGAAAGCA 664
DB 6630 GGTGATGTCCTACACATGAGAGCGGCGCTGTATCAAGCATGCGGTGGAGGAGAAAGCA 6689
QY 665 AGCTGCCATCAACGGGTTGAGCACTTTTGTGCTGATCAACCAACATGATGTACGCTTA 724
```

```
DB 6690 AGCTGCCATCAATGCACTGAGCAACTCTTTGTCTCCGTCACCAACAATTGGTATGCTTA 6749
QY 725 CCAATCTCCGACGCGCAAGCCAGAGCGGACAGAAAGTACTCTTGAACAGACTGCAAAATCC 784
DB 6750 CAACATCTTCGACGCGCAAGCCAGAGCGGACAGAAAGTACTCTTGAACAGACTGCAAGTCC 6809
QY 785 TGAAGATCACTACACGAGACGTCGCTCAAGAGATGAAAGGCGGACAGTTCAGTAAAG 844
DB 6810 TGAAGATCACTACCGGAGACGTCGCTCAAGAGATGAAAGGCGGACAGTTCAGTAAAG 6869
QY 845 CTAACTCTTATAGTGAAGAGAGGCTGCAAGCTGACCGCCCAATTCGAGCTGCAATCTTA 904
DB 6870 CTAACTCTTATAGTGAAGAGAGGCTGTAAGCTGTAAGCGCCCAATTCGAGCTGCAATCTTA 6929
QY 905 AATTGGCTATGAGGCAAGAGACGTCGCGAACCTATTCAGCAAGGCCATTAAACCATCC 964
DB 6930 AATTGGCTATGAGGCAAGAGACGTCGCGAACCTATTCAGCAAGGCCATTAAACCATCC 6989
QY 965 GCTCGGTGGAGAGACTTGTGGAAGACATGAAACCAATTCAGCAAGGCCATTAAACCATCC 1024
DB 6990 GCTCGGTGGAGAGACTTGTGGAAGACATGAAACCAATTCAGCAAGGCCATTAAACCATCC 7049
QY 1025 CAAAAAATGAGGTTTCTGCGTCCAAACGAGAGAGAGGCGGCAAGCCAGCTGCTTA 1084
DB 7050 CAAAAAATGAGGTTTCTGCGTCCAAACGAGAGAGAGGCGGCAAGCCAGCTGCTTA 7109
QY 1085 TCGTGTCCGACACTGGGGGCTCGTGTGTGTCGAGAAATAGGCGCTCTATAGAGTGTCT 1144
DB 7110 TCGTATTCAGATTTGGGGGTTGCTGTGTGTCGAGAAATAGGCGCTCTATAGAGTGTCT 7169
QY 1145 CCACCTCCCTCAGGCTGTGATGAGGACTCTGTGTCAGATTCAGATTCCTGGAACAGC 1204
DB 7170 CCACCTCCCTCAGGCTGTGATGAGGACTCTGTGTCAGATTCAGATTCCTGGAACAGC 7229
QY 1205 GGGTGAATTCCTGAGTGAACGCTTGAATCAAGAGAACCCCTATGAGGCTTTGATATG 1264
DB 7230 GGGTGAATTCCTGAGTGAATTCGTCGAAAGGAGAAATGCGCTATGAGGCTTTGATATG 7289
QY 1265 ACACCCGCTGTTTGAATCAACAGTCACTGAGATGACATCCGCTGTAGAGAGTCAATTT 1324
DB 7290 ACACCCGCTGTTTGAATCAACAGTCACTGAGATGACATCCGCTGTAGAGAGTCAATTT 7349
QY 1325 ATCAATGTTTGAATCTGAGCCCGGAGGACAGGACCATATAGGCTGCTCAAGAGCGGC 1384
DB 7350 ACCAATGTTTGAATCTGAGCCCGGAGGACAGGACCATATAGGCTGCTCAAGAGCGGC 7409
QY 1385 TTTATATGAGGAGTCCCTGATCTAATTCAAAAGGACAGAACTGGCTATGCGCGTGC 1444
DB 7410 TTTATATGAGGAGTCCCTGATCTAATTCAAAAGGACAGAACTGGCTATGCGCGTGC 7469
QY 1445 GCGGAGGAGGCTGTGACGACTAGCTGCGGTAATACCTCAATGTTACTTGAAGGCTT 1504
DB 7470 GCGGAGGAGGCTGTGACGACTAGCTGCGGTAATACCTCAATGTTACTTGAAGGCTT 7529
QY 1505 CTGACAGCTGTGAGCTGCAAGGCTCCAGAGCTGACGATGCTGTATGCGGAGACGAGC 1564
DB 7530 CTGACAGCTGTGAGCTGCAAGGCTCCAGAGCTGACGATGCTGTATGCGGAGACGAGC 7589
QY 1555 TTTGTGTTATGTGAGAGCGCGGAAACCAAGAGAGAGCGGCGAGCTTACGAGTCTTCA 1624
DB 7590 TTTGTGTTATGTGAGAGCGCGGAAACCAAGAGAGAGCGGCGAGCTTACGAGGCTTCA 7649
QY 1635 CCGAGGCTATAGACTTACTCTGCCCCCCCCCGGAGACCGGCCCAACCAATATTCAGCC 1684
DB 7650 CCGAGGCTATAGACTTACTCTGCCCCCCCCCGGAGACCGGCCCAACCAATATTCAGCT 7709
QY 1685 TGAAGTATTAACATCATGCTCTCTCAATGCTGCTGCGGACGATCATCTGCAAAA 1744
DB 7710 TGAAGTATTAACATCATGCTCTCTCAATGCTGCTGCGGACGATCATCTGCAAAA 7769
QY 1745 GGGTATTAACCTCAACCGGTGACCC 1769
```


QY	1085	TCGATTCGCCAGACTTTGGGGGGCGGTGHTGTGGAGAAAATGGCCCTCTATGACGCTGCT	11444
Db	7110	TCGTAATCCCAATTTGGGGGTTCTGTGTGTGGAGAAAATGGCCCTTTACGATGTGGTCT	7169
QY	1145	CCACCCCTCCCTCAGGCTGTGATGGGCTCCTCGTACGATTCGATATTTCTCTGGACAGC	1204
Db	7170	CCACCCCTCCCTCAGGCGGTGATGGGCTCTTCAATACGATTTCCAAATATCTCTCTGGACAGC	72229
QY	1205	GGGTGCAATTCCTGGTGAACGCTGGGAATCAAAAAGACCCCTATGGGCTTTTGATATNG	12645
Db	7230	GGGTGCAATTCCTGGTGAATGCTGGGAAGCGAABAAAATGCCCTATGGGCTTTGCAATNG	7289
QY	1265	ACACCCGCTGTTTGTGACTCAACAGTCACTGAGAAATGATCCGTTAGAGAGTCAATTT	1324
Db	7290	ACACCCGCTGTTTGTGACTCAACAGTCACTGAGAAATGATCCGTTAGAGAGTCAATCT	7349
QY	1335	ATCAATGTTTGTGACTTTGGCCCCCGAAGCGAAGCAGGCGCATAGAGTGTGCTCAAGAGCGGC	1384
Db	7350	ACCAATGTTTGTGACTTTGGCCCCCGAAGCGAAGCAGGCGCATAGAGTGTGCTCAAGAGCGGC	7409
QY	1385	TTTATATGGGGGGTCCCGCTGACTAATTCAAAAGGCGAAGCTGGGGCTATGGCGGGTGC	1444
Db	7410	TTTATATGGGGGGTCCCGCTGACTAATTTTAAAAGGCGAAGCTGGGGCTATGGCGGGTGC	7469
QY	1445	GCAGGAGCGGCGGTGTGACGACTAGCTCGGTAATACCTTCACATGTTACTTGAAGGCT	1504
Db	7470	GCAGGAGCGGCGGTGTGACTAGCAGACTCGGGTAATACCTTCACATGTTACTTGAAGGCT	7529
QY	1505	CTGCAAGCTGTGTGAGCTGCAAAAGCTCCAGAGCTGACGATGTGTGTGGCGAAGCGGC	1564
Db	7530	CTGCGGCTGTGTGAGCTGCAAAAGCTCCAGAGCTGACGATGTGTGTGGCGAAGCGGC	7589
QY	1565	TTGTGCTATCTGTGAGAGCGCGGGAACCCAGAGAGAGCGCGCGAGCTTACGAGTCTTCA	1624
Db	7590	TTGTGCTATCTGTGAGAAAGCGCGGGAACCCAGAGAGAGCGCGCGAGCTTACGAGTCTTCA	7649
QY	1625	CGAGGCTATGACTAGTACTCTGCCCCCGGGGAGCCGCGCCCAACAGAAATACGACC	1684
Db	7650	CGAGGCTATGACTAGTACTCTGCCCCCGGGGAGCCGCGCCCAACAGAAATACGACT	7709
QY	1685	TGGAGTTGATTAACATCATGCTCCTGCATGTGTGGTGGCGCAGATGCAATCTGGCAAAA	1744
Db	7710	TGGAGTTGATTAACATCATGCTCCTGCATGTGTGAGTGTGAGTGGCGCAGATGCAATCTGGCAAAA	7769
QY	1745	GGGTATTAACCTCAACCCGCTGACC 1769	
Db	7770	GGGTATTAACCTCAACCCGCTGACC 7794	
RESULT 6			
AAD25333			
XX	ID	AAD25333 standard; cDNA, 7141 BP.	
XX	AC	AAD25333;	
XX	DT	12-MAR-2002 (first entry)	
XX	DE	Hepatitis C virus (HCV) adaptive replicon 5'NTR-EMCV/HCVrepVII cDNA.	
XX	KM	Hepatitis C virus; HCV; transfection; infection; virus neutralisation;	
XX	KW	gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;	
XX	EMCV/HCVrepVII; 8s.		
XX	OS	Hepatitis C virus.	
XX	PN	MO200189364-A2.	
XX	PD	29-NOV-2001.	
XX	PF	23-MAY-2001; 2001WO-US016822.	
XX	PR	23-MAY-2000; 2000US-00576989.	

Query Match	Best Local Similarity	Matches 1632; Conservative	Score 1552.2; DB 6; Length 7141;	Pred. No. 0;	Mismatches 133; Indels 0; Gaps 0
5	ATGATTTCCCAAGCGCTGATCCTCATTCGAGGCCAACCTCCTGTGGCGCGCATGAGATGGGG	64			
4533	ATGATCTCCCGAGACGCTGATCCTCATTCGAGGCCAACCTCCTGTGGCGCGCATGAGATGGGG	459			
65	GGAACATTACCCGCGCTGAGAGTCAGAGAAACAAGTGTATTCCTGACCTTTTCGACCCGC	124			
4593	GGAACATTACCCGCGCTGAGAGTCAGAGAAACAAGTGTATTCCTGACCTTTTCGACCCGC	465			
125	TCCGAGCGGAGAGATGAGCGGGAGTGTCCGTCCTCCGCGGAGATCTCGCGAAATCCA	184			
4653	TCCGAGCGGAGAGATGAGCGGGAGTGTCCGTCCTCCGCGGAGATCTCGCGAAATCCA	471			
185	AGAAATTTCCCAACGAGCGATGCGCGGATGGGACGCGCGGATTAACAACCTCCGCGTGG	244			
4713	GGAATTTCCCTCGAGCGATGCGCGGATGGGACGCGCGGATTAACAACCTCCGCGTGG	477			
245	AGCTCTGGAAGGCGCGGAGTACGTCCTCCGATGAGTATCATGGGTGCCACTGCCACTTA	304			
4773	AGCTCTGGAAGGCGCGGAGTACGTCCTCCGATGAGTATCATGGGTGCCACTGCCACTTA	483			
305	CTAAGACCCCTTCCTATACCACTTCCAGCGGAAAGAGCAAGTTGTTCTGACAGATCCA	364			
4833	CCAAGGCCCTTCCTGATACCACTTCCAGCGGAAAGAGCAAGTTGTTCTGACAGATCCA	489			
365	CGAGTCTTTCGCGCTGGCGGAGCTGTGCCACAAAGGCTTTTGGTGAAGCTCCGGACGCTGG	424			

CC release, production of adaptive HCV variants capable of more efficiency
CC replication in cell culture, production of HCV variants with altered
CC tissue or species tropism, establishment of alternative animal models for
CC inhibitor evaluation including those supporting HCV variant replication,
CC development of cell-free HCV replication assays, production of
CC immunogenic HCV particles for vaccination, engineering of attenuated HCV
CC derivatives as possible vaccine candidates, engineering of attenuated or
CC defective HCV derivatives for expression of heterologous gene products
CC for gene therapy and vaccine applications and for utilisation of the HCV
CC glycoproteins for targeted delivery of therapeutic agents to the liver
CC or other cell types with appropriate receptors. Vaccine comprising these
CC sequences is useful for inducing immunoprotection to HCV in a primate.
CC The present sequence is Hepatitis C virus (HCV) adaptive replicon
CC HCVrep/NS2-5B CDNA
XX
SQ

Sequence 7789 BP; 1602 A; 2320 C; 2175 G; 1692 T; 0 U; 0 Other;

Query Match 87.7%; Score 1552.2; DB 6; Length 7789;

Best Local Similarity 92.5%; Pred. No. 0;

Matches 1632; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

```
QY 5 ATGACTCCCAAGCGTGAAGCTATCGAGGCAACCTCTGAGGCGATGAGTGGGCG 64
Db ATGACTCCCAAGCGTGAAGCTATCGAGGCAACCTCTGAGGCGATGAGTGGGCG 5240
QY 65 GGGACATTAACCCGCGTGAAGTCAAGAAACAAAGTATCTTGAAGTCTTTCAGACCGC 124
Db GGAACATCAACCCGCGTGAAGTCAAGAAATTAAGTATTTTGAAGTCTTTCAGACCGC 5300
QY 125 TCCGAGCGGAGAGATGAGCGGAAAGTCCGTCGCCGCGAGATCTTCGGAATATCCA 184
Db TCCAGCGGAGAGATGAGGAAAGTATCCGTTCCGCGGAGATCTTCGGAAGTATCCA 5360
QY 185 AGAATTCCTCCACAGAGATGCCGATGGGCAAGCCGGATTAACAACCTCCGCTGCTG 244
Db GGAATTCCTCCACAGAGATGCCGATGGGCAAGCCGGATTAACAACCTCCGCTGCTG 5420
QY 245 AGTCTGGAAGGCGCCGGAATGACGTCCCTCCAGTGGTATGAGTGGTCCCACTGCACTA 304
Db AGTCTGGAAGGCGCCGGAATGACGTCCCTCCAGTGGTATGAGTGGTCCCACTGCACTA 5480
QY 305 CTAAACACCCCTCTTATCAACTCTCAAGGAAAGAGAGAGTGTGTTGACAGAAATCCA 364
Db CTAAGGACCCCTCTTATCAACTCTCAAGGAAAGAGAGAGTGTGTTGACAGAAATCCA 5540
QY 365 CCGTGTCTTCTGAGCGGCGAGCTTGCACAAAGGCTTTGGTATGCTCCGAGCCGTGG 424
Db CCGTGTCTTCTGAGCGGCGAGCTTGCACAAAGGCTTTGGTATGCTCCGAGCCGTGG 5600
QY 425 CCGTGAAGCGGCGAGCGGCAACGCGCCCTCTGACCAATCTTCGACGAGCGGAGAG 484
Db CCGTGAAGCGGCGAGCGGCAACGCGCTCTCTGACCAAGCCCTTCGACGAGCGGAGAG 5660
QY 485 GATTCGACGTTGATGCTATTTCTTCATGAGCCGCTTGAAGGAGAGCGCGGAGAG 544
Db GATTCGACGTTGATGCTATTTCTTCATGAGCCGCTTGAAGGAGAGCGCGGAGAG 5720
QY 545 ATTCGAGCGAGGAGCTTGGTCTACCGTGAAGGAGAGGCGGTAAGAGACGCTGCTCT 604
Db ATTCGAGCGAGGAGCTTGGTCTTACCGTGAAGGAGAGGCTGAAGAGACGCTGCTCT 5780
QY 5721 ATTCGAGCGAGGAGCTTGGTCTTACCGTGAAGGAGAGGCTGAAGAGACGCTGCTCT 5780
QY 605 GCTGATGCTCTTACATGAGAGCGGCTCTGATCAAGCATGCGCTGCGAGAGAAACA 664
Db GCTGATGCTCTTACATGAGAGCGGCTCTGATCAAGCATGCGCTGCGAGAGAAACA 5840
QY 665 AGCTGCCCATCAAGCGGCTTGAAGCACTTTTCTGCTGCTACCAACAATGCTTACGCTA 724
Db AGCTGCCCATCAATGAGCACTTTCCTGCTGCTACCAACAATGCTTACGCTTACGCTA 5900
QY 725 CCACATCCCGGAGCGGAGCGGCGAGAAAGAGTCACTTTTGAAGAGCTGCAAAATCC 784
Db CACATCTCGAGCGGAGCGGCGGAGAGAGGTCACCTTTGAAGAGCTGCAAGGCTCC 5960
```

```
QY 785 TGAAGATCACTACAGAGAGCTGTCTCAAGAGATGAAGAGCGGCTCCACAGTTAAG 844
Db TGAAGATCACTACAGAGAGCTGTCTCAAGAGATGAAGAGCGGCTCCACAGTTAAG 6020
QY 845 CTAACTTCTTATCACTAAGAGAGCTTGAAGCTGACCGCCACATTCGGCAATCTTA 904
Db CTAACTTCTTATCACTAAGAGAGCTTGAAGCTGACCGCCACATTCGGCAATCTTA 6080
QY 905 AATTGGCTAATGGGCAAGGACGCTGGAAACCTATCCAGGAGGCGATTAACCAATCC 964
Db AATTGGCTAATGGGCAAGGACGCTGGAAACCTATCCAGGAGGCGATTAACCAATCC 6140
QY 965 GCTCCGTGAGAGAGCTTGTGGAAGACACTGAACCAATGACACCAATCATATG 1024
Db GCTCCGTGAGAGAGCTTGTGGAAGACACTGAACCAATGACACCAATCATATG 6200
QY 6141 GCTCCGTGAGAGAGCTTGTGGAAGACACTGAAGACCAATGACACCAATCATATG 6200
QY 1025 CAAAAATGAGGTTTCTGCGTCCAAACGAGAGAGAGGCGGCAAGCAGCTGCTTGA 1084
Db CAAAAATGAGGTTTCTGCGTCCAAACGAGAGAGAGGCGGCAAGCAGCTGCTTGA 6260
QY 1085 TCGGTTCCTCAGACTTGGGGGTCGCTGTGTCGAGAAATGAGCCCTATGACGTGCT 1144
Db TCGGTTCCTCAGACTTGGGGGTCGCTGTGTCGAGAAATGAGCCCTATGACGTGCT 6320
QY 1145 CCACCTCTCTCAGGCTGTGAATGGGCTCTCTGTAACGATTCAGTATTCCTCGAGACG 1204
Db CCACCTCTCTCAGGCTGTGAATGGGCTCTCTGTAACGATTCAGTATTCCTCGAGACG 6380
QY 6321 CCACCTCTCTCAGGCTGTGAATGGGCTCTCTGTAACGATTCAGTATTCCTCGAGACG 6380
QY 1205 GGGTGAAGTCTCTGTGTAAGCCTTGAAGATCAAGAGAACCCCTATGAGGCTTTGATATG 1264
Db GGGTGAAGTCTCTGTGTATCTCTGAAGAGCAAGAAATGCTCATGAGGCTTTGATATG 6440
QY 1265 AACCCGCTGTGTGACTGCAACAGTCACTGGAATGACATCCGCTGAGAGAGTCAATTT 1324
Db AACCCGCTGTGTGACTGCAACAGTCACTGGAATGACATCCGCTGAGAGAGTCAATTT 6500
QY 6441 AACCCGCTGTGTGACTGCAACAGTCACTGGAATGACATCCGCTGAGAGAGTCAATTT 6500
QY 1325 ATCAATGTTGACTTGGGCGCCCGAGAGCGAGAGCCATTAAGTCTGCTCAAGAGCGGC 1384
Db ATCAATGTTGACTTGGGCGCCCGAGAGCGAGAGCCATTAAGTCTGCTCAAGAGCGGC 6560
QY 6501 ACCAATGTTGACTTGGGCGCCCGAGAGCGAGAGCCATTAAGTCTGCTCAAGAGCGGC 6560
QY 1385 TTTATATGGGGGTCCTCTGACTAATTTCAAAAGGCGAGAACTGGGCTATGCGGGTCC 1444
Db TTTATATGGGGGTCCTCTGACTAATTTCAAAAGGCGAGAACTGGGCTATGCGGGTCC 6620
QY 6561 TTTATATGGGGGTCCTCTGACTAATTTCAAAAGGCGAGAACTGGGCTATGCGGGTCC 6620
QY 1445 GCGGAGGCGGCTGTGTAAGCACTAGCTGCGGTAATPACCTCACAATGTTAGAGGCT 1504
Db GCGGAGGCGGCTGTGTAAGCACTAGCTGCGGTAATPACCTCACAATGTTAGAGGCT 6680
QY 1505 CTGAGGCTGTGAGCTGCAAGGCTCCAGAGCTGCAAGATGCTGCTGCGAGAGAGCGCC 1564
Db CTGAGGCTGTGAGCTGCAAGGCTCCAGAGCTGCAAGATGCTGCTGCGAGAGAGCGCC 6740
QY 6681 CTGAGGCTGTGAGCTGCAAGGCTCCAGAGCTGCAAGATGCTGCTGCGAGAGAGCGCC 6740
QY 1565 TTTGCTGTTATCTGTGAGAGCGCGGAAACCAAGAGAGAGCGGCGAGGCTTCA 1624
Db TTTGCTGTTATCTGTGAGAGAGCGCGGAAACCAAGAGAGAGCGGCGAGGCTTCA 6800
QY 6741 TTTGCTGTTATCTGTGAGAGAGCGCGGAAACCAAGAGAGAGCGGCGAGGCTTCA 6800
QY 1625 CGGAGGCTATGACTAAGGACTCTGCCCCCGCGGAGAGCGGCGCAACCAAGATACGACC 1684
Db CGGAGGCTATGACTAAGGACTCTGCCCCCGCGGAGAGCGGCGCAACCAAGATACGACC 6860
QY 6801 CGGAGGCTATGACTAAGGACTCTGCCCCCGCGGAGAGCGGCGCAACCAAGATACGACC 6860
QY 1685 TGAAGTGAATTAACATCATGCTCTTCCATATGTGTGCTGCGCAAGATGATCTGGCAAAA 1744
Db TGAAGTGAATTAACATCATGCTCTTCCATATGTGTGCTGCGCAAGATGATCTGGCAAAA 6920
QY 6861 TGAAGTGAATTAACATCATGCTCTTCCATATGTGTGCTGCGCAAGATGATCTGGCAAAA 6920
QY 1745 GGGTATTAATCACTCAACCGGTGACCC 1769
Db GGGTATTAATCACTCAACCGGTGACCC 6945
```

RESULT 8

AAD25329

ID AAD25329 standard; CDNA; 7987 BP.
XX

AC AAD25329;
XX
DT 12-MAR-2002 (first entry)
XX
DE Hepatitis C virus (HCV) adaptive replicon VII cDNA.
XX
KM Hepatitis C virus; HCV; transfection; infection; virus neutralisation;
KM Gene therapy; vaccine; immunoprotection; hepatotropic; viraemic; liver;
KM adaptive replicon VII; ss.
XX
OS Hepatitis C virus.
XX MO200189364-A2.
XX
XX 29-NOV-2001.
XX
XX 23-MAY-2001; 2001MO-US016822.
XX
XX 23-MAY-2000; 2000US-00576989.
XX
XX (UNIM) UNIV WASHINGTON.
XX
XX Rice CM, Blight KJ;
XX
XX MPI; 2002-066755/09.
XX
XX Hepatitis C virus variants having greater transfection efficiency and
PT ability to survive subpassage, useful as a vaccine for immunizing primate
PT to the virus, comprise non-naturally occurring viral sequences.
XX
XX
XX Disclosure; Page 84-87; 174pp; English.
XX
XX The invention relates to Hepatitis C virus (HCV) variants which include
CC polynucleotides comprising non-naturally occurring HCV sequence and HCV
CC variants that have a transfection efficiency and ability to survive
CC subpassage greater than HCV that have wild-type polyprotein coding
CC regions. The polynucleotides of the invention are useful for identifying
CC a cell line that is permissive for infection with HCV and detecting
CC replication of HCV in cells of the cell line. They are also useful for
CC testing a compound for anti-viral properties and for inhibiting HCV
CC infection. They are also useful for the generation of defined HCV virus
CC stocks to develop in vitro and in vivo assays for virus neutralisation,
CC attachment, penetration and entry, structure/function studies on HCV
CC proteins and RNA elements and identification of new antiviral targets, a
CC systematic survey of cell culture systems and conditions to identify
CC those that support wild-type and variant HCV RNA replication and particle
CC release, production of adaptive HCV variants capable of more efficiency
CC replication in cell culture, production of HCV variants with altered
CC tissue or species tropism, establishment of alternative animal models for
CC inhibitor evaluation including those supporting HCV variant replication,
CC development of cell-free HCV replication assays, production of
CC immunogenic HCV particles for vaccination, engineering of attenuated HCV
CC derivatives as possible vaccine candidates, engineering of attenuated or
CC defective HCV derivatives for expression of heterologous gene products
CC for gene therapy and vaccine applications and for utilisation of the HCV
CC glycoproteins for targeted delivery of therapeutic agents to the liver
CC or other cell types with appropriate receptors. Vaccine comprising these
CC sequences is useful for inducing immunoprotection to HCV in a primate.
CC The present sequence is Hepatitis C virus (HCV) adaptive replicon VII
CC cDNA
XX
XX
SQ Sequence 7987 BP; 1647 A; 2368 C; 2242 G; 1730 T; 0 U; 0 Other;

Query Match 87.7%; Score 1552.2; DB 6; Length 7987;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1632; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

QY 5 ATGACTCCAGAGCGTGAAGTCAATGAGGCAACCTCCGTGGCGGCAATGAGGCG 64
DB 5381 ATGACTCCAGAGCGTGAAGTCAATGAGGCAACCTCCGTGGCGGCAATGAGGCG 5440
QY 65 GGGACATTAACCGCGTGAAGTCAAGAAAGAGTAATCTGAGCTTTTGAGCCGCG 124
DB 65 GGGACATTAACCGCGTGAAGTCAAGAAAGAGTAATCTGAGCTTTTGAGCCGCG 124

DB 5441 GGAACATCACCGCGTGAAGTCAAGAAATGAAGTAATTTTGAAGCTTTTGAGCCGCG 5500
QY 125 TCCGAGCGAGAGAGATGAGCGGAAAGTGTCCGTCCCGGAGATCTTCGGAAATCCA 184
DB 5501 TCCAGCGGAGAGAGATGAGAGGAAAGTCCGTCCCGGAGATCTTCGGAGGTCGA 5560
QY 185 AGAAATTTCCACAGAGAGATGCCGCAATGGGACAGCCCGGATTCACACCTCCGTGTGG 244
DB 5561 GGAATTTCCCTCGAGAGATGCCATATGGGACAGCCCGGATTCACACCTCCGTGTGG 5620
QY 245 AGTCTGGAAGAGCCCGGAGATGAGTCCCTCAAGTGTATCAATGGTCCCATCTCACTTA 304
DB 5621 AGTCTGGAAGAGCCCGGAGATGAGTCCCTCAAGTGTATCAATGGTCCCATCTCACTTA 5680
QY 305 CTAAAGACCCCTCTTAATCACTTCACAGAGAAAGAGACAGTTGTTCTGACGAATCCA 364
DB 5681 CCAAGGCCCTCTCGAATCACTTCACAGAGAAAGAGACAGTTGTTCTGACGAATCTTA 5740
QY 365 CCGTGTCTTCTGCGCTGGGAGAGCTTGGCCACAAAGGCTTTGGTAGCTCCGACCGTGG 424
DB 5741 CCGTGTCTTCTGCGCTGGGAGAGCTTGGCCACAAAGGCTTTGGTAGCTCCGACCGTGG 5800
QY 425 CCGTGCACAGCGGACAGGCAACCGCCCTCTGAGCAATCTTCGACAGACGCGAGAGCAG 484
DB 5801 CCGTGCACAGCGGACAGGCAACCGCCCTCTGAGCAATCTTCGACAGACGCGAGAGCAG 5860
QY 485 GATCTGACGTTGATGTTGATTTCTTCATGATGCCCCCTTGAAGGGGAGACCGGAGACCCG 544
DB 5861 GATCTGACGTTGATGTTGATTTCTTCATGATGCCCCCTTGAAGGGGAGACCGGAGACCCG 5920
QY 545 ATCTGACGAGCGGAGTCTGAGTCTACCGTGAAGAGAGCGGAGAGAGCGTGTGCTGCT 604
DB 5921 ATCTGACGAGCGGAGTCTGAGTCTACCGTGAAGAGAGCGGAGAGAGCGTGTGCTGCT 5980
QY 605 GCTGAGTGTCTTACATGATGAGACAGGCGCTGTGATCAAGCATGCGTGGAGAGAAAGCA 664
DB 5981 GCTGAGTGTCTTACATGATGAGACAGGCGCTGTGATCAAGCATGCGTGGAGAGAAAGCA 6640
QY 665 AGCTGCCATCAACGCGTTGAGCACTTTTGTGCTGATCAACCAATGTTGATGCTTA 724
DB 6041 AGCTGCCATCAACGCGTTGAGCACTTTTGTGCTGATCAACCAATGTTGATGCTTA 7100
QY 725 CCAATCCCGGAGCGGAGCGGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 784
DB 6101 CCAATCCCGGAGCGGAGCGGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6160
QY 785 TGAGACATCACTACAG 844
DB 6161 TGAGACATCACTACAG 6220
QY 845 CTAAAGCTTTATCAATGAG 904
DB 6221 CTAAAGCTTTATCAATGAG 6280
QY 905 AATTTGGCATGAGGAG 964
DB 6281 AATTTGGCATGAGGAG 6340
QY 965 GCTCCGTTGGAG 1024
DB 6341 GCTCCGTTGGAG 6400
QY 1025 CAAAAAATGAGGTTTCTGCTGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1084
DB 6401 CAAAAAATGAGGTTTCTGCTGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6460
QY 1085 TGGTGTTCAGAGCTTGGGAG 1144
DB 6461 TGGTGTTCAGAGCTTGGGAG 6520
QY 1145 CCAACCTCCCTCAGAGCTGTGATGAGGCTCTCTGATAGAGATTCAGATATCTCTCGAGACAG 1204
DB 6521 CCAACCTCCCTCAGAGCTGTGATGAGGCTCTCTGATAGAGATTCAGATATCTCTCGAGACAG 6580


```
QY 1205 GGGTCAGTTCCTGCGTGAACGCGCTGGAATCAAGAGACCCTATGCGCTTGGCATATG 1264
    |||||
DB 6581 GGGTCAGTTCCTGCGTGAATGCTGGAAGCAAGAAAGCCCTATGCGCTTGGCATATG 6640
QY 1265 ACACCCGCTGTTTGACTCAACAGTCACTGAGAAATGACATCCGTGTAGAGAGTCAATT 1324
    |||||
DB 6641 ACACCCGCTGTTTGACTCAACAGTCACTGAGAAATGACATCCGTGTAGAGAGTCAATT 6700
QY 1325 ATCAATGTTGACTGTTGGCCCCGGAAGCAGACGGCCATAAGTGTGCTCAGAGAGCGCC 1384
    |||||
DB 6701 ACCAATGTTGACTGTTGGCCCCGGAAGCAGACGGCCATAAGTGTGCTCAGAGAGCGCC 6760
QY 1385 TTTATATCGGGGGGCTCCCTGACTAAATTCAAAAGGAGACTGCGGCTATCGCCGGTGC 1444
    |||||
DB 6761 TTTATATCGGGGGGCTCCCTGACTAAATTCAAAAGGAGACTGCGGCTATCGCCGGTGC 6820
QY 1445 GCGCGAGCGCGCTGCTGACGACTAGTGGGTAAATCCCTCAATGTTACTTGAAGGCT 1504
    |||||
DB 6821 GCGCGAGCGCGTGTACTGACGACGAGCTGCGTAAATCCCTCAATGTTACTTGAAGGCGG 6880
QY 1505 CTGCAAGCTGTGAGCTGCAAGCTCCAGAGCTGCAAGCTGCTGCTGCGGAGACGCGC 1564
    |||||
DB 6881 CTGCGGCTGTGAGCTGCAAGCTCCAGAGCTGCAAGCTGCTGCTGCGGAGACGAGC 6940
QY 1565 TTGCTGTTATCTGTGAGAGCGCGGAAACCAGAGAGACGCGGCTACGAGTCTTCA 1624
    |||||
DB 6941 TTGCTGTTATCTGTGAAAGCGGGGAGCCCAAGAGAGACGAGGCTACGAGGCTTCA 7000
QY 1625 CGGAGGCTATGACTAGTACTCTGCCCCCGGGGAGCCGCCCAACCAAGAAATGAGC 1684
    |||||
DB 7001 CGGAGGCTATGACTAGTACTCTGCCCCCGGGGAGCCGCCCAACCAAGAAATGAGC 7060
QY 1685 TGGAGTGATACATCATGCTCTCCATGATGTCGGTGGCGAGCATGATCGGCAAA 1744
    |||||
DB 7061 TGGAGTGATACATCATGCTCTCCATGATGTCGGTGGCGAGCATGATCGGCAAA 7120
QY 1745 GGGTATACTACTCAACCCGTGAGCC 1769
    |||||
DB 7121 GGGTATACTACTCAACCCGTGAGCC 7145

RESULT 9
AAD25321
ID AAD25321 standard; cDNA; 7987 BP.
XX
AC AAD25321;
XX
DT 12-MAR-2002 (first entry)
XX
DE Hepatitis C virus (HCV) repIDBartman/delta2U's cDNA.
XX
KM Hepatitis C virus; HCV; transfection; infection; virus neutralisation;
KM gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;
KM ss.
XX
OS Hepatitis C virus.
XX
PN WO200189364-A2.
XX
PD 29-NOV-2001.
XX
PF 23-MAY-2001; 2001WO-US016822.
XX
PR 23-MAY-2000; 2000US-00576989.
XX
PA (UNITW ) UNITV WASHINGTON.
XX
PI Rice CM, Blight KJ;
XX
DR WPI; 2002-066755/09.
XX
PT Hepatitis C virus variants having greater transfection efficiency and
```

```
PT ability to survive subpassage, useful as a vaccine for immunizing primate
PT to the virus, comprise non-naturally occurring viral sequences.
XX
PS Claim 44; Page 66-69; 174pp; English.
XX
CC The invention relates to Hepatitis C virus (HCV) variants which include
CC polynucleotides comprising non-naturally occurring HCV sequence and HCV
CC variants that have a transfection efficiency and ability to survive
CC subpassage greater than HCV that have wild-type polypeptide coding
CC regions. The polynucleotides of the invention are useful for identifying
CC a cell line that is permissive for infection with HCV and detecting
CC replication of HCV in cells of the cell line. They are also useful for
CC testing a compound for anti-viral properties and for inhibiting HCV
CC infection. They are also useful for the generation of defined HCV virus
CC stocks to develop in vitro and in vivo assays for virus neutralisation,
CC attachment, penetration and entry, structure/function studies on HCV
CC proteins and RNA elements and identification of new antiviral targets, a
CC systematic survey of cell culture systems and conditions to identify
CC those that support wild-type and variant HCV RNA replication and particle
CC release, production of adaptive HCV variants capable of more efficiency
CC replication in cell culture, production of HCV variants with altered
CC tissue or species tropism, establishment of alternative animal models for
CC inhibitor evaluation including those supporting HCV variant replication,
CC development of cell-free HCV replication assays, production of
CC immunogenic HCV particles for vaccination, engineering of attenuated HCV
CC derivatives as possible vaccine candidates, engineering of attenuated or
CC defective HCV derivatives for expression of heterologous gene products
CC for gene therapy and vaccine applications and for utilisation of the HCV
CC glycoproteins for targeted delivery of therapeutic agents to the liver
CC or other cell types with appropriate receptors. Vaccine comprising these
CC sequences is useful for inducing immunoprotection to HCV in a primate.
CC The present sequence is Hepatitis C virus (HCV) repIDBartman/delta2U's
CC cDNA
SQ Sequence 7987 BP; 1647 A; 2368 C; 2243 G; 1729 T; 0 U; 0 Other;
XX
Query Match 87.7%; Score 1552.2; DB 6; Length 7987;
Beet Local Similarity 92.5%; Pred. No. 0;
Matches 1632; Conservative 0; Mismatches 133; Indels 0; Gaps 0;
QY 5 ATGACTTCCGAGAGGCTTAACCTCATCGAGGCAACCTCTGTGCGGAGATGATGGCG 64
    |||||
DB 5381 ATGACTCCCGAGAGCTGACCTCATCGAGGCAACCTCTGTGCGGAGAGATGGCG 5440
QY 65 GGGACATTACCCTGCGTGAATCAGAGAAACAAGTATGTAATCTGTGACTTTTCAGACCGC 124
    |||||
DB 5441 GGAACATACCCCGCTGAGATCAAGAAATAAGTAATTTGGACTTTTCAGACCGC 5500
QY 125 TCCGAGCGGAGAGATGAGCGGAAAGTGTCCGTCCGCGGAGATCTGCGGAATCCA 184
    |||||
DB 5501 TCCAAAGCGGAGAGATGAGAGGAAATATCCGTTCCGGCGGAGATCTGCGGAGATCCA 5560
QY 185 AGAATTTCCACACAGCATGCCCCGATGAGGACCGCCGGATTAACAACCTTCGCTGCTG 244
    |||||
DB 5561 GGAATTTCCCTCGAGCGATGCCCCATATGGGACCGCCGGATTAACAACCTTCACGTTAG 5620
QY 245 AGTCTGGAAGGCCCGGACTACGTCCTCCAGTGTATGATGGGACCACTGCGACCTA 304
    |||||
DB 5621 AGTCTGGAAGAGACCCGAGCTACGTCCTCCAGTGTATGATGGGACCACTGCGACCTG 5680
QY 305 CTAAGACCCCTCTAATACCACTCCACGAGAAAGAGACAGTTGTTGACAGAAATCCA 364
    |||||
DB 5681 CCAAGGCCCTCCGATACCACTCCACGAGAAAGAGACAGTTGTTGACAGAAATCCA 5740
QY 365 CCGTGTCTTTCGCTGCGGAGGCTTGTGCAAAAGCTTTTGTGATGCTCCGAGCCTGTGG 424
    |||||
DB 5741 CCGTGTCTTTCGCTGCGGAGGCTTGTGCAAAAGCTTTGAGGAGCTCCGAATGTGTGG 5800
QY 425 CCGTGAAGGAGGAGCAACCGCCCTCTCGAACCAATCTCCGAGACGCGGAGAG 484
    |||||
DB 5801 CCGTGAAGGAGGAGCAACCGCCCTCTCTGACCAACCTCCGAGACGCGGAGAG 5860
QY 485 GATCGACGTTGAGTGTATTCCTTCATGATCCGCCCTTGAAGGAGGAGCCGGGAGACCCCG 544
```



```

Db      5861 GATCCGACGTTGAGTGTACTCTCTCCATGCCCCCTTGAAGGGGAGACCGGGGATCCCG 5920
Qy      545 ATTCAGCGACGGGCTTGGTCTACCGTGAATGAGAGCGGTGAGAGAGCTGCTGCT 604
Db      5921 ATCTCAGCGACGGGCTTGGTCTACCGTGAAGCGAGAGGCTAGTAGAGAGTGTGCTGCT 5980
Qy      605 GCTCGAGTGTCTTACATGAGAGAGGGCTCTGTATCAAGCCATGCGGTGCGAGAGAAAGCA 664
Db      5981 GCTCGAGTGTCTTACATGAGAGAGGGCTCTGTATCAAGCCATGCGGTGCGAGAGAAAGCA 6040
Qy      665 AGCTGCCATCAACGCGGTTGAGCAACTTTGTGCTGCTACCAACAATGCTTACGCTA 724
Db      6041 AGCTGCCATCAATGAGAGAGGCACTTTGTGCTGCTACCAACAATGCTTACGCTA 6100
Qy      725 CCAATATCCGCGAGCGGACGAGCGGCGAGAAAGGTCACTTTTGAACAATGCCAAATCC 784
Db      6101 CAACATCTCGACAGCGGAGCTGCGGCGAGAAAGGTCACTTTTGAACAATGCCAGGTCC 6160
Qy      785 TGGACGATCACTTCCAGAGAGTGTCTCAAGAGATGAAAGGCGTCCACAGTTAAGG 844
Db      6161 TGGACGACCACTTCCGAGAGTGTCTCAAGAGATGAAAGGCGTCCACAGTTAAGG 6220
Qy      845 CTAAAGCTTCTATCAGTAGAGAGAGCTGCAAGCTGACGCCCCACATTTGGGCGCAATCTA 904
Db      6221 CTAAATCTTCTATCCGTGAGAGAGAGCTGTAAGCTGAGCGCCCCACATTTGGGCGCAATCTA 6280
Qy      905 AATTGGCTATGAGGCGCAAGAGAGCTGCGAAGCTATTCAGCAAGGCGCATTTAACCATATCC 964
Db      6281 AATTGGCTATGAGGCGCAAGAGAGCTGCGAAGCTATTCAGCAAGGCGCATTTAACCATATCC 6340
Qy      965 GCTCCGCTGTGGAGAGACTTGTGTGAAGACACTGAACACCAATTGACACCAATCATATGG 1024
Db      6341 GCTCCGCTGTGGAGAGACTTGTGTGAAGACACTGAAGACACCAATTGACACCAATCATATGG 6400
Qy      1025 CAAAAAATGAGGTTTTCTGCGTCAACAGAGAGAGAGGCGGCAAGCGAGCTCGGCTTA 1084
Db      6401 CAAAAAATGAGGTTTTCTGCGTCAACAGAGAGAGAGGCGGCGGCAAGCGAGCTCGGCTTA 6460
Qy      1085 TGTGTGTCCAGAGCTTGGGGGTCGTGTGTGAGAGAAATGAGCGCTCTATGAGTGGTCT 1144
Db      6461 TGTGTATCCAGAGTTTGGGGGTTCTGTGTGTGAGAGAAATGAGCGCTCTATGAGTGGTCT 6520
Qy      1145 CCACCTCTCTCTGAGGCTGTGATGGGCTCTCTGTAAGGATTCAGATATTCTCTGAGACAGC 1204
Db      6521 CCACCTCTCTCTGAGGCTGTGATGGGCTCTCTGTAAGGATTCAGATATTCTCTGAGACAGC 6580
Qy      1205 GGGTCGAGTTCTGCTGTGAACGCTGGAATCAAGAAAGACCCCTATGAGGCTTTGCAATATG 1264
Db      6581 GGGTCGAGTTCTGCTGTGAATGCTGGAAGAGCAAGAAATGCCCTATGAGGCTTTGCAATATG 6640
Qy      1265 ACAACCGCTGTTTGACTCAACAGTCACTGAGATGACATCCGTGTAGAGAGTCAATTT 1324
Db      6641 ACAACCGCTGTTTGACTCAACAGTCACTGAGATGACATCCGTGTAGAGAGTCAATCT 6700
Qy      1325 ATCAATGTTGTGACTTGGCCCCGGAAGCCAGACAGGCGCATTAAGTGTCTACAGAGCGGC 1384
Db      6701 ACCAAATGTTGTGACTTGGCCCCGGAAGCCAGACAGGCGCATTAAGTGTCTACAGAGCGGC 6760
Qy      1385 TTTATATCGGGGGGTCCTGACTAATTCAAAAGGCGAGACCTGCGGTATCGCGGATGCC 1444
Db      6761 TTTATATCGGGGGGTCCTGACTAATTCAAAAGGCGAGACCTGCGGTATCGCGGATGCC 6820
Qy      1445 GCGCGAGCGCGGTGCTGACGACTAGCTGCGGTAAATACCTTCACATGTTACTTGAAGGCTT 1504
Db      6821 GCGCGAGCGCGGTGCTGACGACTAGCTGCGGTAAATACCTTCACATGTTACTTGAAGGCTT 6880
Qy      1505 CTGACACCTGTGAGAGTGCAGAAAGCTCCAGAGCTGACAGATGCTGTGTGGGAGAGCGGC 1564
Db      6881 CTGCGGCTGTGAGAGTGCAGAAAGCTCCAGAGCTGACAGATGCTGTGTGGGAGAGCGGC 6940
Qy      1565 TTTCGTTATCTGTGAGAGCGCGGGAACCGAGAGAGCGGGGAGGCTTACGAGTCTTCA 1624

```

```

Db      6941 TTGTGTTATCTGTGAAGCGCGGGAGCCCAAGAGACGAGCGGAGCTACGGGCTTCA 7000
Qy      1625 CGAGGCTATGACTAGTACTCTGCCCCCGGGGAGACCCGCGCCCAACAGAAATAGAAC 1684
Db      7001 CGAGGCTATGACTAGTACTCTGCCCCCGGGGAGACCCGCGCCCAACAGAAATAGAAC 7060
Qy      1685 TGGAGTTGATTAACATCATGCTCTCCATATGTGTGCGTCCGCGACATGATCTGGCAAAA 1744
Db      7061 TGGAGTTGATTAACATCATGCTCTCCATATGTGTGAGTCCGCGACATGATCTGGCAAAA 7120
Qy      1745 GGGTATTAATCACTACCCGCTGAGCCC 1769
Db      7121 GGGTATTAATCACTACCCGCTGAGCCC 7145

RESULT 10
AAD25324
ID AAD25324 standard; cDNA; 7987 BP.
XX
AC AAD25324;
XX
DT 12-MAR-2002 (first entry)
XX
DE Hepatitis C virus (HCV) adaptive replicon VI cDNA mutant.
XX
KW Hepatitis C virus; HCV; transfection; infection; virus neutralisation;
KW gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;
KW adaptive replicon VI; mutant; ss.
OS Hepatitis C virus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1601..7758
FT FT /*tag= a
FT FT /product= "HCV adaptive replicon VI protein"
FT FT /note= "CDS does not include both start and stop codon"
FT FT /partial
FT FT replace(536, G)
FT FT /*tag= c
XX
EN WO200189364-A2.
XX
PD 29-NOV-2001.
XX
PF 23-MAY-2001; 2001WO-US016822.
XX
PR 23-MAY-2000; 2000US-00576989.
XX
PA (UNIM ) UNIV WASHINGTON.
XX
PA PI
XX
XX Rice CM, Blight RJ;
XX
DR WPI; 2002-066755/09.
DR P-PSDB; AAB15720, AAB15721.
XX
PT Hepatitis C virus variants having greater transfection efficiency and
PT ability to survive subpassage, useful as a vaccine for immunizing primate
PT to the virus, comprise non-naturally occurring viral sequences.
XX
XX disclosure; Page 74-77; 174dp; English.
XX
XX The invention relates to Hepatitis C virus (HCV) variants which include
XX polynucleotides comprising non-naturally occurring HCV sequence and HCV
XX variants that have a transfection efficiency and ability to survive
XX subpassage greater than HCV that have wild-type polypeptide coding
XX regions. The polynucleotides of the invention are useful for identifying
XX a cell line that is permissive for infection with HCV and detecting
XX replication of HCV in cells of the cell line. They are also useful for
XX testing a compound for anti-viral properties and for inhibiting HCV

```

infection. They are also useful for the generation of defined HCV virus stocks to develop in vitro and in vivo assays for virus neutralisation, attachment, penetration and entry, structure/function studies on HCV proteins and RNA elements and identification of new antiviral targets, a systematic survey of cell culture systems and conditions to identify those that support wild-type and variant HCV RNA replication and particle release, production of adaptive HCV variants capable of more efficiency replication in cell culture, production of HCV variants with altered tissue or species tropism, establishment of alternative animal models for inhibitor evaluation including those supporting HCV variant replication, development of cell-free HCV replication assays, production of immunogenic HCV particles for vaccination, engineering of attenuated HCV derivatives as possible vaccine candidates, engineering of attenuated or defective HCV derivatives for expression of heterologous gene products for gene therapy and vaccine applications and for utilisation of the HCV glycoproteins for targeted delivery of therapeutic agents to the liver or other cell types with appropriate receptors. Vaccine comprising these sequences is useful for inducing immunoprotection to HCV in a primate. The present sequence is Hepatitis C virus (HCV) adaptive replicon VI cDNA mutant. This sequence is generated by the mutation g to t at position 5336 of HCVrep1bBarman/Availi cDNA

Sequence 7987 BP; 1646 A; 2368 C; 2243 G; 1730 T; 0 U; 0 Other;

Query Match 87.7%; Score 1552.2; DB 6; Length 7987;

Best Local Similarity 92.5%; Pred. No. 0; Matches 1632; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

```
QY 5 ATGACTTCCCAACGCTGACCTCATCGAGGCCAACCTCTGTGGGGGAGTGAATGGGCG 64
DB 5381 ATGACTCCCGGAGCGTGCATCATCGAGGCCAACCTCTGTGGGGGAGTGAATGGGCG 5440
QY 65 GGGACATTACCCGCGTGAAGTGAAGAAACAAGTATTCCTGACCTTTTCGACCCGCG 124
DB 5441 GGAATCATACCCGCGTGAAGTGAAGAAACAAGTATTCCTGACCTTTTCGACCCGCG 5500
QY 125 TCCGACCGGAGAGATGAGCGGAAAGTGTCCGTCGGCGAGATCTTCGCGAAATCCA 184
DB 5501 TCCAGACGGAGAGATGAGCGGAAAGTGTCCGTCGGCGAGATCTTCGCGAGATCCA 5560
QY 185 AGAAATTTCCCAACGAGATGAGCGGAAAGTGTCCGTCGGCGAGATCTTCGCGAAATCCA 244
DB 5561 GGAATTTCCCAACGAGATGAGCGGAAAGTGTCCGTCGGCGAGATCTTCGCGAGATCCA 5620
QY 245 AGTCTCGAAGGCGCCGAGCTACGTCCTCCAGTGGTATCAGGTCGCCACTGCCACCTA 304
DB 5621 AGTCTCGAAGGCGCCGAGCTACGTCCTCCAGTGGTATCAGGTCGCCACTGCCACTG 5680
QY 305 CTAAAGACCCCTCTATACCACTCCACGGAAGAAGAGCAAGTTGTTCTGACAGATCCA 364
DB 5681 CCAAGGCCCTCTCGATACCACTCCACGGAAGAAGAGCAAGTTGTTCTGACAGATCCA 5740
QY 365 CCGTGTCTTCGCGCTGGGGAGCTTGGCCAAAGGCTTTTGTAGCTCCGACCGTGG 424
DB 5741 CCGTGTCTTCGCGCTGGGGAGCTTGGCCAAAGGCTTTTGTAGCTCCGACCGTGG 5800
QY 425 CCGTGAAGGCGGAGCGGCAACCGCCCTCTGACCAATCTCCGACGAGCGGAGAG 484
DB 5801 CCGTGAAGGCGGAGCGGCAACCGCCCTCTGACCAATCTCCGACGAGCGGAGAG 5860
QY 485 GATCTGAGCTGATGCTGATTCCTTCATGAGCGCCCTTTGAGGGGAGAGCGGAGAGCCCG 544
DB 5861 GATCTGAGCTGATGCTGATTCCTTCATGAGCGCCCTTTGAGGGGAGAGCGGAGAGCCCG 5920
QY 545 ATCTAGAGAGCGGGCTTGGTCTTACCGTGAAGAGAGCGGCTGAGAGCGTGTCTGCT 604
DB 5921 ATCTAGAGAGCGGGCTTGGTCTTACCGTGAAGAGAGCGGCTGAGAGCGTGTCTGCT 5980
QY 605 GCTCATGCTCTACACATGAGCAGGCGCTCTGATCAACGCGCATGCGCTGGGAGAGAA 664
DB 5981 GCTCATGCTCTACACATGAGCAGGCGCTCTGATCAACGCGCATGCGCTGGGAGAGAA 6040
QY 665 AGCTGCCCATCAACGCGTTGAGCAACTCTTTGCTGGTCAACCAACATGCTCTACGCTA 724
```

```
DB 6041 AGCTGCCCATCAATGACATGAGCAACTCTTTGCTCCGTCACACAACTTGGTCTATGCTA 6100
QY 725 CCACATCCCGGAGCGCAACCGGAGGAGAGAGTCACTTTGACAGACTGCAAAATCC 784
DB 6101 CAACATTCCTGAGCGCAACCGGAGGAGAGAGTCACTTTGACAGACTGCAAGTCC 6160
QY 785 TGAGCATCACTACAGGAGCGTGTCAAGAGATGAAGGAGCGTCCACAGTTAAG 844
DB 6161 TGAGCATCACTACAGGAGCGTGTCAAGAGATGAAGGAGCGTCCACAGTTAAG 6220
QY 845 CTAAGCTTCTATCAATGAGAGAGAGCTTGAAGCTGACGCGCCCAATTCGAGCAATCTA 904
DB 6221 CTAAGCTTCTATCAATGAGAGAGAGCTTGAAGCTGACGCGCCCAATTCGAGCAATCTA 6280
QY 905 AATTGGCTATGAGGAGAGAGAGCTTGAAGCTGACGCGCCCAATTCGAGCAATCTA 964
DB 6281 AATTGGCTATGAGGAGAGAGAGCTTGAAGCTGACGCGCCCAATTCGAGCAATCTA 6340
QY 965 GCTCCGTGGAGAGACTTGTGGAAGACATGAACCAATTAACACACATCATATG 1024
DB 6341 GCTCCGTGGAGAGACTTGTGGAAGACATGAACCAATTAACACACATCATATG 6400
QY 1025 CAAAAATGAGGTTTCTGCGTCCAAACGAGAGAGAGCGGCAAGCCAGCTGCTT 1084
DB 6401 CAAAAATGAGGTTTCTGCGTCCAAACGAGAGAGAGCGGCAAGCCAGCTGCTT 6460
QY 1085 TCGTGTCCGAGACTTGGGGGCTCGTGTGTCGAGAAATGGCCCTCTATACAGTGTCT 1144
DB 6461 TCGTGTCCGAGACTTGGGGGCTCGTGTGTCGAGAAATGGCCCTCTATACAGTGTCT 6520
QY 1145 CCACCTCCCTCAGGCTTGAATGGGCTCTGTGAGATTCAGATATCTCTCGAGACGC 1204
DB 6521 CCACCTCCCTCAGGCTTGAATGGGCTCTGTGAGATTCAGATATCTCTCGAGACGC 6580
QY 1205 GGGTGAATTCCTGAGTGAACGCTGGAATTAAGAGAGCCCTATAGGCTTTGCAATAG 1264
DB 6581 GGGTGAATTCCTGAGTGAACGCTGGAATTAAGAGAGCCCTATAGGCTTTGCAATAG 6640
QY 1265 ACACCCGCTGTTTGAACCAACAGTCACTGAGATGACATCGGTGAGAGATCAATTT 1324
DB 6641 ACACCCGCTGTTTGAACCAACAGTCACTGAGATGACATCGGTGAGAGATCAATTT 6700
QY 1325 ATCAATGTTGATCTGGGCGCCCGGAGGCGAGAGCCATTAAGTGTCTCAAGAGCGGC 1384
DB 6701 ATCAATGTTGATCTGGGCGCCCGGAGGCGAGAGCCATTAAGTGTCTCAAGAGCGGC 6760
QY 1385 TTTATATCGGGGCTCCCTGATCTAATTCAAAAGGAGAGAGCTGAGCTTATCGCGTGC 1444
DB 6761 TTTATATCGGGGCTCCCTGATCTAATTCAAAAGGAGAGAGCTGAGCTTATCGCGTGC 6820
QY 1445 GCGGAGGCGGCTGACGACCTAGCTGCGGCTTATACCTTCACATGTTAATTGAAGGCT 1504
DB 6821 GCGGAGGCGGCTGACGACCTAGCTGCGGCTTATACCTTCACATGTTAATTGAAGGCT 6880
QY 1505 CTGACGCTGTGAGAGCTCAAGCTCCAGAGCTGACGATGCTGTGTCGAGAGAGCGGC 1564
DB 6881 CTGACGCTGTGAGAGCTCAAGCTCCAGAGCTGACGATGCTGTGTCGAGAGAGCGGC 6940
QY 1565 TTTGTGTTATCTGAGAGCGCGGAGACCAAGAGAGAGCGGCGAGAGCTTACGAGTCTTCA 1624
DB 6941 TTTGTGTTATCTGAGAGCGCGGAGACCAAGAGAGAGCGGCGAGAGCTTACGAGTCTTCA 7000
QY 1625 CGAGGCTATATCAATGATCTCTCCCGCCCGGAGAGAGCGGCGCAACCAATTAAGAGC 1684
DB 7001 CGAGGCTATATCAATGATCTCTCCCGCCCGGAGAGAGCGGCGCAACCAATTAAGAGC 7060
QY 1685 TGAGTTGATTAACATCATGCTCTCTCAATGTGTGCGTGCAGAGATCATCTGGCAAAA 1744
DB 7061 TGAGTTGATTAACATCATGCTCTCTCAATGTGTGCGTGCAGAGATCATCTGGCAAAA 7120
QY 1745 GGGTATATCACTGACCCGAGAGCC 1769
```



```

Db      6641 ACACCCGCTGTTTGACTCAAGCGTCACTAGAAATGACATCCGATTGTTGAGATCAATCT 6700
Qy      1325 ATCAATGTTGTACTTGGCCCCCGAAGCAGAGCCATAAGTGTGCTCACAGAGCGGC 1384
Db      6701 ACCAATGTTGTACTTGGCCCCCGAAGCAGAGCCATAAGTGTGCTCACAGAGCGGC 6760
Qy      1385 TTTATATCGGGGGTCCCTGACTTAATTCAAAAGGCGAAGTGGCGCTATCGCCGGTGCC 1444
Db      6761 TTTACATCGGGGGCCCCCTGACTTAATTCAAAAGGCGAAGTGGCGCTATCGCCGGTGCC 6820
Qy      1445 GGGCGAGCGCGCTGCTGAGAGACTAGTGGGGTAATTCCTTCACATGTTACTTGAAGGCTT 1504
Db      6821 GGGCGAGCGGGTACTGAGACAGCTGGGTGAATTAACCTTCACATGTTACTTGAAGGCGCG 6880
Qy      1505 CTGCAAGCTGTGAGCTGCAAGCTTCAGAGACTGCAAGATGCTCGTGTGCGAGACGCGC 1564
Db      6881 CTGCGGCTGTGAGCTGCAAGCTTCAGAGACTGCAAGATGCTCGTGTGCGAGACGAGC 6940
Qy      1565 TTGTCGTATCTGTGAGAGCGCGGGAACCCAGAGAGACGCGGCGAGCTTACAGATCTTCA 1624
Db      6941 TTGTCGTATCTGTGAAAGCGCGGGGACCAAGAGAGACGAGCGGCTTACAGGCTTCA 7000
Qy      1625 CGGAGGCTATGACTAGTACTCTGCCCCCCCCCGGGGACCCGCCCAACCAAGATTCGACC 1684
Db      7001 CGGAGGCTATGACTAGTACTCTGCCCCCCCCCGGGGACCCGCCCAACCAAGATTCGACT 7060
Qy      1685 TGGAGTTGATACATCATATGCTCTCCCAATGTGTGCGGTGCGAGACATGATCGGCAAAA 1744
Db      7061 TGGAGTTGATACATCATATGCTCTCCCAATGTGTGCGGTGCGAGACATGATCGGCAAAA 7120
Qy      1745 GGGTACTACTCACTCAACCCCGTGAACC 1769
Db      7121 GGGTACTACTCACTCAACCCCGTGAACC 7145

```

RESULT 12

AAD25322 standard; cDNA; 7989 BP.

AAD25322;

12-MAR-2002. (first entry)

Hepatitis C virus (HCV) rep/Bartman/Availi cDNA.

Hepatitis C virus; HCV; transfection; infection; virus neutralisation;

gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;

86.

Hepatitis C virus.

Location/Qualifiers

Key CDS 1801..7758

FT misc_feature /tag= a /product= "HCVrep/Bartman polyprotein"

FT /tag= b /note= "Nucleotide creating Availi site"

FN WO200189364-A2.

PD 29-NOV-2001.

PP 23-MAY-2001; 2001WO-US016822.

PR 23-MAY-2000; 2000US-00576989.

PA (UNIM) UNIV WASHINGTON.

PI Rice CM, Blight KJ;

DR WPI; 2002-06755/09.

P-PSDB; AAB15717.

XX Hepatitis C virus variants having greater transfection efficiency and
PT ability to survive subpassage, useful as a vaccine for immunizing primate
PT to the virus, comprise non-naturally occurring viral sequences.

PS Claim 44; Page 69-71; 174p; English.

XX The invention relates to Hepatitis C virus (HCV) variants which include
CC polynucleotides comprising non-naturally occurring HCV sequence and HCV
CC variants that have a transfection efficiency and ability to survive
CC subpassage greater than HCV that have wild-type polypeptide coding
CC regions. The polynucleotides of the invention are useful for identifying
CC a cell line that is permissive for infection with HCV and detecting
CC replication of HCV in cells of the cell line. They are also useful for
CC testing a compound for anti-viral properties and for inhibiting HCV
CC infection. They are also useful for the generation of defined HCV virus
CC stocks to develop in vitro and in vivo assays for virus neutralisation,
CC attachment, penetration and entry, structure/function studies on HCV
CC proteins and RNA elements and identification of new antiviral targets, a
CC systematic survey of cell culture systems and conditions to identify
CC those that support wild-type and variant HCV RNA replication and particle
CC release, production of adaptive HCV variants capable of more efficiency
CC replication in cell culture, production of HCV variants with altered
CC tissue or species tropism, establishment of alternative animal models for
CC inhibitor evaluation including those supporting HCV variant replication,
CC development of cell-free HCV replication assays, production of
CC immunogenic HCV particles for vaccination, engineering of attenuated HCV
CC derivatives as possible vaccine candidates, engineering of attenuated or
CC defective HCV derivatives for expression of heterologous gene products
CC for gene therapy and vaccine applications and for utilisation of the HCV
CC glycoproteins for targeted delivery of therapeutic agents to the liver
CC or other cell types with appropriate receptors. Vaccine comprising these
CC sequences is useful for inducing immunoprotection to HCV in a primate.
CC The present sequence is Hepatitis C virus (HCV) rep/Bartman/Availi cDNA

XX Sequence 7989 BP; 1647 A; 2369 C; 2242 G; 1731 T; 0 U; 0 Other;

Query Match 87.7%; Score 1552.2; DB 6; Length 7989;

Best Local Similarity 92.5%; Pred. No. 0;

Matches 1632; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

```

Qy      5 ATGACTTCCAGACGCTGACCTCATCGAGCCCAACTCTGTGGCGGAGATGAGGCG 64
Db      5381 ATGACTTCCCGAGAGCTGACCTCATCGAGCCCAACTCTGTGGCGGAGATGAGGCG 5440
Qy      65 GGGACATACCCGGGTGAGTCAGAGAACAGAGTAATCTGGACTTTTCAGCCGCG 124
Db      5441 GGAACATACCCGGGTGAGTCAGAGAAATTAAGTAATTTGGACTTTTCAGCCGCG 5500
Qy      125 TCCGAGCGAGAGAGATGAGCGGAAAGTCCGTCGCCGCGAGATCCTGCGGAATCCA 184
Db      5501 TCCAGCGGAGAGAGATGAGCGGAAAGTCCGTCGCCGCGAGATCCTGCGGAATCCA 5560
Qy      185 AGAAATTCACACGAGATGCGCCGATGCGGACCCCGGATTAACAACCTTCGCTGTGG 244
Db      5561 GGAATTCCTCGAGCGATGCGCCATATGGCGACCCCGGATTAACAACCTTCACGTTAG 5620
Qy      245 AGTCTGGAAGCGCCCGGACTAGCTCCCTCCAGTGTGTAATGAGTGGCCACTGCACTTA 304
Db      5621 AGTCTGGAAGCGCCCGGACTAGCTCCCTCCAGTGTGTAATGAGTGGCCACTGCGGCTG 5680
Qy      305 CTAAGACCCCTCTATACCACTTCAAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 364
Db      5681 CCAAGGCCCTCTCGATACCACTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5740
Qy      365 CCGTGTCTTCTGCGCTGCGAGCTTGCACAAAGGCTTTGTAGCTTCGAGACGCTGCG 424
Db      5741 CCGTGTCTTCTGCGCTTGGCGAGCTCGCCACAAAGAGCTTGCAGAGCTCCGAATGTGCG 5800
Qy      425 CCGTGTGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 484
Db      5801 CCGTGTGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5860

```

QY 485 GATCTGACGTTGATGCTGATCTCTCCATGAGCCCCCTTGAGGGGGAGGCCGGGGAGCCCGG 544
 DB 5861 GATCCGACGTTGATGCTGATCTCTCCATGAGCCCCCTTGAGGGGGAGGCCGGGGAGCCCGG 5920
 QY 545 ATCTCAGCAGCGGGCTTGGTCTACCGTGAAGAGAGCGGGTGAAGAGCTGCTGCT 604
 DB 5921 ATCTCAGCAGCGGGCTTGGTCTACCGTGAAGAGAGCGGGTGAAGAGCTGCTGCTGCT 5980
 QY 605 GCTCCATGCTCTACATGAGCAGAGCGCTCTGATCAACGCCATCCGCTGGGAGAGAAAGCA 664
 DB 5981 GCTCCATGCTCTACATGAGCAGAGCGCTCTGATCAACGCCATCCGCTGGGAGAGAAAGCA 6040
 QY 665 AGCTGCCATCAACCGGTTGAGCACTCTTGTGCTGATCAACCAATGCTGCTAGCTA 724
 DB 6041 AGCTGCCATCAACCGGTTGAGCACTCTTGTGCTGATCAACCAATGCTGCTAGCTA 6100
 QY 725 CCACATCCCGCAGCGAGCCAGCGCAGAGAGAGAGTCACTTTGACAGACTGCAATCC 784
 DB 6101 CAACATCTCGCAGCGAGCCAGCGCAGAGAGAGTCACTTTGACAGACTGCAATCC 6160
 QY 785 TGAACGATCACTACAGAGCTGCTCAAGAGATGAAGGGGAAAGCGCTCCAGTTAAG 844
 DB 6161 TGAACGATCACTACAGAGCTGCTCAAGAGATGAAGGGGAAAGCGCTCCAGTTAAG 6220
 QY 845 CTAGCTTCTATCAGTGAAGAGAGCTGAGCTGAGCGGCCCACTGGGCCCAATCTA 904
 DB 6221 CTAGCTTCTATCAGTGAAGAGAGCTGAGCTGAGCGGCCCACTGGGCCCAATCTA 6280
 QY 905 AATTGGCTATGAGGGGAGAGAGAGCTGCGAACCCTATCCAGCAGAGCCATTAACCAATCC 964
 DB 6281 AATTGGCTATGAGGGGAGAGAGAGCTGCGAACCCTATCCAGCAGAGCCATTAACCAATCC 6340
 QY 965 GCTCCGCTGAGAGAGCTGTTGGAAGACATGGAACCAATGAACCAACCATCATG 1024
 DB 6341 GCTCCGCTGAGAGAGCTGTTGGAAGACATGGAACCAATGAACCAACCATCATG 6400
 QY 1025 CAAAAAATAGAGTTTCTGCTCAACGAGAGAGAGCGGCAAGCAGCTGCGCTTA 1084
 DB 6401 CAAAAAATAGAGTTTCTGCTCAACGAGAGAGAGCGGCAAGCAGCTGCGCTTA 6460
 QY 1085 TCGTGTTCACAGACTTGGGGGCTCCGTGTGTCAGAGAAATGGCCCTCTATGAGCTGCT 1144
 DB 6461 TCGTGTTCACAGACTTGGGGGCTCCGTGTGTCAGAGAAATGGCCCTCTATGAGCTGCT 6520
 QY 1145 CCAACCTCCCTCAGGCTGTGATGAGGCTCTCTGTAAGGATTCAGATTCCTCTGAGACG 1204
 DB 6521 CCAACCTCCCTCAGGCTGTGATGAGGCTCTCTGTAAGGATTCAGATTCCTCTGAGACG 6580
 QY 1205 GGGTGAAGTTCCTGAGTGAAGCGCTGGAATCAAGAGAGCCCTATGAGGCTTGCATG 1264
 DB 6581 GGGTGAAGTTCCTGAGTGAAGCGCTGGAATCAAGAGAGCCCTATGAGGCTTGCATG 6640
 QY 1265 ACACCCGCTGTTTGACTCAACAGTCACTGAGATGACATCCGTGATGAGAGATCAATT 1324
 DB 6641 ACACCCGCTGTTTGACTCAACAGTCACTGAGATGACATCCGTGATGAGAGATCAATT 6700
 QY 1325 ATCAATGTTTGACTTGGCCCCCGAGAGCAGACAGGCGCATAGGCTGCTCAAGAGCGCG 1384
 DB 6701 ACCAATGTTTGACTTGGCCCCCGAGAGCAGACAGGCGCATAGGCTGCTCAAGAGCGCG 6760
 QY 1385 TTTTATATGGGGGTCCCTGAGCTAATTTCAAAAGGGAGAACTGGCGCTATCGCGGTGCC 1444
 DB 6761 TTTTATATGGGGGTCCCTGAGCTAATTTCAAAAGGGAGAACTGGCGCTATCGCGGTGCC 6820
 QY 1445 GGGCGAGCGGCTGCTGAGCAGTACTGCGGTATATCCCTCAATTTACTTTGAAGGCTT 1504
 DB 6821 GGGCGAGCGGCTGCTGAGCAGTACTGCGGTATATCCCTCAATTTACTTTGAAGGCTT 6880
 QY 1505 CTGACAGCTGTGAGCTGCAAAAGCTCAAGAGCTGCAAGATGCTGCTGTCGAGAGAGCGCC 1564
 DB 6881 CTGACAGCTGTGAGCTGCAAAAGCTCAAGAGCTGCAAGATGCTGCTGTCGAGAGAGCGCC 6940
 QY 1565 TTGTCTTATCTGTGAGAGCGGGAGAACCAAGAGAGACCGGGAGACCTTACAGATCTTCA 1624

DB 6941 TTGTCTTATCTGTGAAGCGGGAGACCCAGAGAGACAGAGGAGCTTACGGGCTTCA 7000
 QY 1625 CGAGGCTATGACTAGTACTTGTGCCCCCGGGAGACCGGCCCAACAGAAATAGAC 1684
 DB 7001 CGAGGCTATGACTAGTACTTGTGCCCCCGGGAGACCGGCCCAACAGAAATAGAC 7060
 QY 1685 TGAAGTTGATACATGATGCTCTCAATGATCGGTGCGGAGCAGATGATCGGCAAA 1744
 DB 7061 TGAAGTTGATACATGATGCTCTCAATGATCGGTGCGGAGCAGATGATCGGCAAA 7120
 QY 1745 GGGTATCTACTCACTCAACCGGTGACCC 1769
 DB 7121 GGGTATCTACTCACTCAACCGGTGACCC 7145
 RESULT 13
 AAD25326
 ID AAD25326 standard; cDNA; 7989 BP.
 AC AAD25326;
 DT 12-MAR-2002 (first entry)
 XX
 DE Hepatitis C virus (HCV) adaptive replicon V mutant cDNA.
 XX
 KW Hepatitis C virus; HCV; transfection; infection; virus neutralisation;
 KW gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;
 KW adaptive replicon V; mutant; ss.
 XX
 OS Hepatitis C virus.
 OS Synthetic.
 XX
 FH Key
 FT CDS
 FT /tag= a
 FT /product= "NS5A variant of HCV adaptive replicon V"
 FT /note= "CDS does not include both start and stop codon"
 FT /partial
 FT replace(5320, G)
 FT mutation
 FT /tag= b
 XX
 PN MO200189364-A2.
 XX
 PD 29-NOV-2001.
 XX
 PF 23-MAY-2001; 2001WO-US016822.
 XX
 PR 23-MAY-2000; 2000US-00576989.
 XX
 PA (UNITW) UNIV WASHINGTON.
 XX
 PI Rice CM, Blight KJ;
 XX
 DR WPI: 2002-066755/09.
 DR P-PSDB; AAE15724.
 XX
 PT Hepatitis C virus variants having greater transfection efficiency and
 PT ability to survive subpassage, useful as a vaccine for immunizing primate
 PT to the virus, comprise non-naturally occurring viral sequences.
 XX
 PS Disclosure; Page 80-83; 174pp; English.
 XX
 CC The invention relates to Hepatitis C virus (HCV) variants which include
 CC polynucleotides comprising non-naturally occurring HCV sequence and HCV
 CC variants that have a transfection efficiency and ability to survive
 CC subpassage greater than HCV that have wild-type polypeptide coding
 CC regions. The polynucleotides of the invention are useful for identifying
 CC a cell line that is permissive for infection with HCV and detecting
 CC replication of HCV in cells of the cell line. They are also useful for
 CC testing a compound for anti-viral properties and for inhibiting HCV
 CC infection. They are also useful for the generation of defined HCV virus
 CC stocks to develop in vitro and in vivo assays for virus neutralisation,

RESULT 14
AAD25325
ID AAD25325 standard; cDNA; 7989 BP.
XX
AC AAD25325;
XX
DT 12-MAR-2002 (first entry)
XX
XX Hepatitis C virus (HCV) adaptive replicon II cDNA mutant.
XX
XX Hepatitis C virus; HCV; transfection; infection; virus neutralisation;
XX gene therapy; vaccine; immunoprotection; hepatotropic; viraemic; liver;
XX adaptive replicon II; mutant; ss.
XX
OS Hepatitis C virus.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH 1801..7758
FT CDS
FT CDS /tag= a
FT CDS /product= "HCV adaptive replicon II protein"
FT CDS /tag= b
FT CDS /product= "NS5A protein of HCV adaptive replicon II"
FT CDS /note= "CDS does not include both start and stop codon"
FT CDS /partial
XX
XX MO200189364A2.
XX
XX 29-NOV-2001.
XX
XX 23-MAY-2001; 2001MO-US016822.
XX
XX 23-MAY-2000; 2000US-00576989.
XX
XX (UNIM) UNIV WASHINGTON.
XX
XX Rice CM, Blight RJ;
XX
XX MPI; 2002-06755/09.
XX P-PSDB; AAB15722, AAB15723.
XX
XX Hepatitis C virus variants having greater transfection efficiency and
XX ability to survive subpassage, useful as a vaccine for immunising primate
XX to the virus, comprise non-naturally occurring viral sequences.
XX
XX Example 1; Page 77-80; 174pp; English.
XX
XX The invention relates to Hepatitis C virus (HCV) variants which include
XX polynucleotides comprising non-naturally occurring HCV sequence and HCV
XX variants that have a transfection efficiency and ability to survive
XX subpassage greater than HCV that have wild-type polypeptide coding
XX regions. The polynucleotides of the invention are useful for identifying
XX a cell line that is permissive for infection with HCV and detecting
XX replication of HCV in cells of the cell line. They are also useful for
XX testing a compound for anti-viral properties and for inhibiting HCV
XX infection. They are also useful for the generation of defined HCV virus
XX stocks to develop in vitro and in vivo assays for virus neutralisation,
XX attachment, penetration and entry, structure/function studies on HCV
XX proteins and RNA elements and identification of new antiviral targets, a
XX systematic survey of cell culture systems and conditions to identify
XX those that support wild-type and variant HCV RNA replication and particle
XX release, production of adaptive HCV variants capable of more efficiency
XX replication in cell culture, production of HCV variants with altered
XX tissue or species tropism, establishment of alternative animal models for
XX inhibitor evaluation including those supporting HCV variant replication,
XX development of cell-free HCV replication assays, production of
XX immunogenic HCV particles for vaccination, engineering of attenuated HCV
XX derivatives as possible vaccine candidates, engineering of attenuated or
XX defective HCV derivatives for expression of heterologous gene products
XX for gene therapy and vaccine applications and for utilisation of the HCV
XX glycoproteins for targeted delivery of therapeutic agents to the liver

CC or other cell types with appropriate receptors. Vaccine comprising these
CC sequences is useful for inducing immunoprotection to HCV in a primate.
CC The present sequence is Hepatitis C virus (HCV) adaptive replicon II cDNA
CC mutant. This sequence is generated by the mutation a to g at position
CC 5289 of HCVrep1bBartman/Avail cDNA
XX
SQ Sequence 7989 BP; 1644 A; 2369 C; 2245 G; 1731 T; 0 U; 0 Other;
Query Match 87.7%; Score 1552.2; DB 6; Length 7989;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1632; Conservative 0; Mismatches 133; Indels 0; Gaps 0;
QY 5 ATGACTTCCGAGCGGCTGACCTCATGAGGCGCAACCTCGTGGCGGATGAGTGGCG 64
DB 5381 ATGACTTCCGAGCGGCTGACCTCATGAGGCGCAACCTCGTGGCGGAGTGGCG 5440
QY 65 GGGACATTAACCCGCGTGAAGTCAAGAACAGGTAGTAATCTGTGACTTTTCGACCGC 124
DB 5441 GGAACATTAACCCGCGTGAAGTCAAGAACAGGTAGTAATTTTGAATCTTTTCGACCGC 5500
QY 125 TCCGAGCGGAGAGAGATGAGCGGAAAGTGTCCGTCGGCGGAGATCTCGGAAATCCA 184
DB 5501 TCCAGCGGAGAGAGATGAGCGGAAAGTGTCCGTCGGCGGAGATCTCGGAGGTCCA 5560
QY 185 AGAAATTCACACGAGGATGCGGATGAGGCGAGCGCGGATTAACAACCTCGCTGTGG 244
DB 5561 GGAATTCCTCCAGGAGTGGCCATATGAGGCGAGCGCGGATTAACAACCTCGCTGTAG 5620
QY 245 AGTCTGGAAGAGCGCGGAGCTACGTCCTTCAGTGTGATGAGGCGGAGCTCACTCA 304
DB 5621 AGTCTGGAAGAGCGCGGAGCTACGTCCTTCAGTGTGATGAGGCGGAGCTCACTCA 5680
QY 305 CTAAAGACCCCTCTTATACCACTTCAAGGAAAGAGACAGTGTTCGACGAATCCA 364
DB 5681 CCAAGGCGGCTCGATACCACTTCAAGGAAAGAGACAGTGTTCGACGAATCCA 5740
QY 365 CGGTGCTTCTGCGCGGAGGCTTGCACAAAGGCTTTGGTACTCCGCGGTCCG 424
DB 5741 CGGTGCTTCTGCGCGGAGGCTTGCACAAAGGCTTTGGTACTCCGCGGTCCG 5800
QY 425 CCGTCGACAGCGGACGAGCAACCGGCTCTGACCAATCTTCGACGAGCGGAGAG 484
DB 5801 CCGTCGACAGCGGACGAGCAACCGGCTCTGACCAAGGCTTCGACGAGCGGAGAG 5860
QY 485 GATCTGACGTTGATGTCGTAATCTTCATGAGGCGGCGGAGGAGCGGAGCGG 544
DB 5861 GATCTGACGTTGATGTCGTAATCTTCATGAGGCGGCGGAGGAGCGGAGCGG 5920
QY 545 ATCTGACGAGCGGCTTGTGCTTACCGTGAAGAGGCGGTAAGAGCTGTCTGCT 604
DB 5921 ATCTGACGAGCGGCTTGTGCTTACCGTGAAGAGGCGGTAAGAGCTGTCTGCT 5980
QY 605 GGTGAGTCTTACATGAGAGAGGCGCTGTGATCAGCGCATGCGTGGAGAGAAAGCA 664
DB 5981 GGTGAGTCTTACATGAGAGAGGCGCTGTGATCAGCGCATGCGTGGAGAGAAAGCA 6040
QY 665 AGTGCCTCAACAGCGGTTGAGCACTTTTGTGCTGTCACCAACATGTTACGCTA 724
DB 6041 AGTGCCTCAACAGCGGTTGAGCACTTTTGTGCTGTCACCAACATGTTACGCTA 6100
QY 725 CCAACATCCGAGCGGAGCGGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 784
DB 6101 CCAACATCCGAGCGGAGCGGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6160
QY 785 TGGACATCACTACAG 844
DB 6161 TGGACATCACTACAG 6220
QY 845 CTAAAGCTTATATGATGAG 904
DB 6221 CTAAAGCTTATATGATGAG 6280
QY 905 AATTGGCTATGAGGAG 964


```
Db      6281 AATTGGCTATGGGGCAAAAGACGTCCGAACTATCCAGCAAGGCCGTTAACACATCC 6340
Qy      965 GGTCCGTGGGAGGAGCTTTGGTGAAGACATGAAACACCAATTGACACACATCATAG 1024
Db      6341 GCTCCGTGGAGGAGCTTGTGTAAGACATGAGCACCAATTGACACACATCATATG 6400
Qy      1025 CAAAAAATGAGGTTTCTGCTCAACAGAGAGAGAGGCGCAAGCAGCTCGCTTA 1084
Db      6401 CAAAAAATGAGGTTTCTGCTCAACAGAGAGAGGCGCGCAAGCAGCTCGCTTA 6460
Qy      1085 TGTGTTCCCAAGCTTTGGGGGTCCTGTGTGCGAGAAAATGGCCCTTATGACGTGCT 1144
Db      6461 TGTATTTCCCAATTTGGGGGTCCTGTGTGCGAGAAAATGGCCCTTATGATGTGCT 6520
Qy      1145 CCACCTCCCTCAGAGCTGATGGGCTCCTCGTACGAGATTCCAGTATTCCTCCGACAC 1204
Db      6521 CCACCTCCCTCAGAGCTGATGGGCTCCTCATACGGAATTCGAATCTCTCCGACAC 6580
Qy      1205 GGGTGAAGTTCCTGTGAACGCTGGAATCAAAAGAGACCCCTATGGGCTTTGCAATG 1264
Db      6581 GGGTGAAGTTCCTGTGAATGCTGGAAGCGAAGAAATGCCCTATGGGCTTCGATATG 6640
Qy      1265 ACACCCGCTGTTTGACTCAACAGTCACTGAGATGACATCCGTGTAGAGAGTCAATT 1324
Db      6641 ACACCCGCTGTTTGACTCAACAGTCACTGAGATGACATCCGTGTAGAGAGTCAATCT 6700
Qy      1325 ATCAATGTTGACTGTTGGCCCCCGAAGCAGACGCGCTAAGTGTGCTCAAGAGCGGC 1384
Db      6701 ACCAATGTTGACTGTTGGCCCCCGAAGCAGACGCGCTAAGTGTGCTCAAGAGCGGC 6760
Qy      1385 TTTATATCGGGGGGCTCCCTGACTAATTCAAAAGGAGAACTGGGCTATCGCCGGTGC 1444
Db      6761 TTTATATCGGGGGGCTCCCTGACTAANTCTAAAGGAGAACTGGGCTATCGCCGGTGC 6820
Qy      1445 GCGCGAGCGGCTGCTGACGACTAGTGGGTAATCCTTCACATGTTACTTGAAGGCT 1504
Db      6821 GCGCGAGCGGCTGCTGACGACTAGTGGGTAATCCTTCACATGTTACTTGAAGGCT 6880
Qy      1505 CTGACGCTGTGAGCTGCAAGGCTCCAGACGTCGACGATGCTCGTGGCGGAGACGCGC 1564
Db      6881 CTGCGGCTGTGAGCTGCAAGGCTCCAGACGTCGACGATGCTCGTGGCGGAGACGACC 6940
Qy      1565 TTGTGTTATCTGTGAGAGCGCGGGAACCCAGAGAGACGCGGCTACGAGTCTTCA 1624
Db      6941 TTGTGTTATCTGTGAAAGCGCGGGAACCCAGAGAGACGCGGCTACGAGGCTTCA 7000
Qy      1625 CCGAGGCTATGACTAGTACTCTGCCCCCGGGGAGCCCGCCCAACCAAGATACGACT 1684
Db      7001 CCGAGGCTATGACTAGTACTCTGCCCCCGGGGAGCCCGCCCAACCAAGATACGACT 7060
Qy      1685 TCGAGTTGTAACATCATGCTCTCCCAATGTGTGCGTGGCGAGCATGTGCGAATA 1744
Db      7061 TCGAGTTGTAACATCATGCTCTCCCAATGTGTGCGTGGCGAGCATGTGCGAATA 7120
Qy      1745 GGGTATTAATCACTCAACCCGTGAACC 1769
Db      7121 GGGTATTAATCACTCAACCCGTGAACC 7145

RESULT 15
ADD93722
ID      ADD93722 standard; DNA; 7989 BP.
XX
AC      ADD93722;
XX
DT      29-JAN-2004 (first entry)
XX
DB      Hepatitis C virus strain H77 (BB7-F1) replicon.
XX
KM      HCV; vaccine; virucide; ss.
OS      Hepatitis C virus.
```

```
XX      WO2003085084-A2.
PN      16-OCT-2003.
XX
PD      03-APR-2003; 2003WO-US010177.
XX
PF      03-APR-2002; 2002US-0369685P.
XX
PR      (SMK) SMITHLINE BECHAM CORP.
PA
XX      Gates A, Gu B, Sarisky RT;
PI      WPI; 2003-804301/75.
DR
XX
PT      New hepatitis C virus (HCV) sub-genomic replicon, useful for facilitating
PT      screening or testing of anti-HCV drugs, comprises a nucleic acid
PT      construct encoding chimeric HCV non-structural proteins, and an NS5B
PT      polymerase gene.
PS
XX      Claim 16; Page 27-32; 159pp; English.
XX
CC      The present sequence comprises a replicating hepatitis C virus (HCV) H77
CC      (BB7-F1) sub-genomic replicon. The invention provides sub-genomic
CC      replicons of HCV comprising a nucleic acid construct encoding chimeric
CC      HCV nonstructural protein and an NS5B polymerase gene. A preferred
CC      replicon comprises an NS3 nucleotide sequence ADD93721 that encodes the
CC      first 75 contiguous N-terminal amino acids of the NS3 of genotype 1b, of
CC      a BB7 strain. A chimeric replicon may comprise an NS3 sequence from any
CC      of the 6 major HCV genotypes and subtypes but has its first 225
CC      nucleotides of the coding sequence replaced by the BB7 strain NS3
CC      sequence, especially where the replicon is from HCV genotype 1a (H77
CC      strain) or genotype 1b (J4 strain). Stable cell lines expressing and
CC      replicating functional replicons containing sequences from HCV genotype
CC      1a (strain H77) or genotype 1b (strain J4) within the prototype 1b
CC      replicon backbone from HCV strain BB7 are provided. These can be used to
CC      screen for compounds that modulate viral replication. The sub-genomic HCV
CC      replicon systems of the invention may provide the foundation for
CC      generating HCV replicons of all 6 major genotypes and subtypes to
CC      facilitate screening, testing and evaluating anti-infective agents for
CC      HCV disease(s).
XX
SQ      Sequence 7989 BP; 1626 A; 2373 C; 2235 G; 1755 T; 0 U; 0 Other;
XX
Query Match      87.7%; Score 1552.2; DB 10; Length 7989;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1632; Conservative 0; Mismatches 133; Indels 0; Gaps 0;
Qy      5 ATGACTTCCAGACGCTGACCTCATCGAGGCCAACCTCTGTGGCGGATGATGGGCG 64
Db      5381 ATGACTCCCGGAGCGCTGACTCATCGAGGCCAACCTCTGTGGCGGAGATGGGCG 5440
Qy      65 GGGACATTAACCCGCTGAGATGACAGAAACAGATGTAATCTTGAATCTTTCCACCCGC 124
Db      5441 GGAACATCAACCCGCTGAGATGACAGAAATGATGTAATTTGAACTTTTCAGACCGC 5500
Qy      125 TCCGAGCGGAGAGATGAGCGGGAAGTGTCCGTCCGCGGAGATCTTGCGGAATCA 184
Db      5501 TCCAGCGGAGAGATGAGCGGGAAGTGTCCGTCCGCGGAGATCTTGCGGAGATCA 5560
Qy      185 AGAAATTCACACAGCGATGCGCGATGGGACGCGCGGATTAACAACCTCCGTGCTGG 244
Db      5561 GGAATTTCCCTCGAGCGATGCGCATATGGGACGCGCGGATTAACAACCTCCACGTGTA 5620
Qy      245 AGTCTTGAAGAGCCCGGACTAAGTCCCTCCAGTGTGATGATGGGTGCCACTGCACTTA 304
Db      5621 AGTCTTGAAGAGACCCGAGTACAGTCCCTCCAGTGTGATGATGGGTGCCACTGCGCTG 5680
Qy      305 CTAAGACCCCTCTATACCACTCCAGAGAGAAAGAGACAGTTGTTGACAGAAATCA 364
Db      5681 CTAAGACCCCTCTATACCACTCCAGAGAGAAAGAGACAGTTGTTGTTGACAGAAATCA 5740
Qy      365 CCGTGTCTTTCGCTCGCGAGGCTTGGCACAAAGGCTTTTGTAGCTCCGAGACCGTGG 424
```


Db 5741 CCGTGTCTTCTGCTTGGCGAGCTGCGCAAGACCTTGGCAGCTCCGAATCGTCGG 5800
Qy 425 CCGTCGACAGCGGACGCGGACCGCCCTCTGACCAATCTCCGACGAGCGGAGCAG 484
Db 5801 CCGTCGACAGCGGACGCGGACCGCCCTCTGACCAAGCTCCGACGAGCGGAGCGCG 5860
Qy 485 GATCTGACGTTGAGTGTATCTCTCAATGCCCCCTTGAAGGGGAGCCGGGGAGCCCG 544
Db 5861 GATCCACGCTTGTGATCTCTCCATGCCCCCTTGAAGGGGAGCCGGGGAGCTCCCG 5920
Qy 545 ATCTGACGACGGGCTTGTGATCTGATGAGGAGCGGGTGAAGAGTGTCTGCT 604
Db 5921 ATCTGACGACGGGCTTGTGATCTGATGAGGAGCGGGTGAAGAGTGTCTGCTGCT 5980
Qy 605 GCTCGATGCTCAACATGAGCAGCGCTCTGATCAAGCCATGCGTGGGAGAGAAC 664
Db 5981 GCTCGATGCTCAACATGAGCAGCGCTCTGATCAAGCCATGCGTGGGAGAGAAC 6040
Qy 665 AGCTGCGCATCAACGCGTGTGAGCACTCTTGTGCTGCTCAACCAATGTCTAGCTTA 724
Db 6041 AGCTGCGCATCAACGCGTGTGAGCACTCTTGTGCTGCTCAACCAATGTCTAGCTTA 6100
Qy 725 CCAATCCCGACGCGGACGCGGAGAGAGGTCACCTTGAAGAGCTGCAATCC 784
Db 6101 CCAATCTCGACGCGGACGCGGAGAGAGGTCACCTTGAAGAGCTGCAATCC 6160
Qy 785 TGGACGATGACTACAGAGCTGTCTCAAGAGATGAGAGGCGTCCACAGTTAAG 844
Db 6161 TGGACGATGACTACAGAGCTGTCTCAAGAGATGAGAGGCGTCCACAGTTAAG 6220
Qy 845 CTGAGCTTCTATCATGAGAGAGAGCTGCAAGCTGACGCCCACTTCGGCCAAATCTA 904
Db 6221 CTGAACTTCTATCGTGTGAGAGAGCTGTGAGCTGACGCCCACTTCGGCCAAATCTA 6280
Qy 905 AATTTGGCTATGGGGGAGAGAGAGTCCGGAACCTATCCAGAGGCGCATTAACACATCC 964
Db 6281 AATTTGGCTATGGGGGAGAGAGAGTCCGGAACCTATCCAGAGGCGCATTAACACATCC 6340
Qy 965 GCTCCGCTGTGAGAGAGCTGTGTGAGAGACATGAAACCCAAATTTGACACCATCATAG 1024
Db 6341 GCTCCGCTGTGAGAGAGCTGTGTGAGAGACATGAAACCCAAATTTGACACCATCATAG 6400
Qy 1025 CAAGAAATGAGTTTCTGCTGTCAACAGAGAGAGAGGCGGAGCCAGCTCGCTTA 1084
Db 6401 CAAGAAATGAGTTTCTGCTGTCAACAGAGAGAGAGGCGGAGCCAGCTCGCTTA 6460
Qy 1085 TGGTGTCCAGAGCTGGGGGCTCGGTGTGCGAGAGAAATGGGCCCTATAGCTGTGCT 1144
Db 6461 TGGTATTTCCAGATTTGGGGGCTGTGTGTGCGAGAGAAATGGGCCCTTATAGATGTGCT 6520
Qy 1145 CCAACCTCTCTCAGGCTGTGATGGGCTCTGTACGAGATTCCAGTATTTCTCTGACAGC 1204
Db 6521 CCAACCTCTCTCAGGCTGTGATGGGCTCTGTACGAGATTCCAGTATTTCTCTGACAGC 6580
Qy 1205 GGGTCGAGTTCTGTGTGAACGCTCGAATCAAGAGAGAGCCCTATGGGCTTTGCATATG 1264
Db 6581 GGGTCGAGTTCTGTGTGAATGCTGTGAAGAGAGAGAGAGCCCTATGGGCTTTGCATATG 6640
Qy 1265 ACAACCGGCTTTTGACTGAACAGTCACTGAGATGATCCGTTGAGAGAGAGCAATTT 1324
Db 6641 ACAACCGGCTTTTGACTGAACAGTCACTGAGATGATCCGTTGAGAGAGAGCAATTT 6700
Qy 1325 ATCAATGTTGTGACTTGTGCCCCCGAAGCAGACAGGCGCATAGGTCTGCTACAGAGCGC 1384
Db 6701 ACCAATGTTGTGACTTGTGCCCCCGAAGCAGACAGGCGCATAGGTCTGCTACAGAGCGC 6760
Qy 1385 TTTATATCGGGGGTCCCCGTGACTTAATCAAGAGGAGAGAGCTGGGCTATCGCGGTGCC 1444
Db 6761 TTTATATCGGGGGTCCCCGTGACTTAATCAAGAGGAGAGAGCTGGGCTATCGCGGTGCC 6820
Qy 1445 GCGGAGCGGCTGTGAGAGACTAGCTGGGTAATACCTCACATGTTACTTGAAGGCT 1504

Db 6821 GCGGAGCGGCTGTGAGAGAGCTGCGGTAATACCTCACATGTTAATTGAAGCGC 6880
Qy 1505 CTGACGCTGTGTGAGCTGCAAGCTCCAGAGCTGCAAGATGCTGTGTGTGAGAGAGCGC 1564
Db 6881 CTGCGGCTGTGTGAGCTGCGAAGCTCCAGAGCTGCAAGATGCTGTGTGTGAGAGAGCGC 6940
Qy 1565 TTGTGTTATCTGTGAGAGCGGGAGAACCGAGAGAGAGCGGGAGCTTACAGATCTTCA 1624
Db 6941 TTGTGTTATCTGTGAGAGCGGGAGAACCGAGAGAGAGCGGGAGCTTACAGATCTTCA 7000
Qy 1625 CCGAGGCTATGACTAGTACTTGTGCCCCCGGGAGACCGGCCCAACGAGATTAAGACC 1684
Db 7001 CCGAGGCTATGACTAGTACTTGTGCCCCCGGGAGACCGGCCCAACGAGATTAAGACC 7060
Qy 1685 TGGAGTTGATTAATCATATGCTCTCCAAATGTGTGCTGCGGACGATGCTTGGCAAAA 1744
Db 7061 TGGAGTTGATTAATCATATGCTCTCCAAATGTGTGCTGCGGACGATGCTTGGCAAAA 7120
Qy 1745 GGGTATTAACCTACCGCGTGAACC 1769
Db 7121 GGGTATTAATCTACCGCGTGAACC 7145

Search completed: February 19, 2005, 01:26:41
Job time : 995.878 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2005, 00:21:46 ; Search time 304.59 Seconds

(without alignments)
9508.540 Million cell updates/sec

Title: US-09-664-363-3

Perfect score: 1770
Sequence: 1 CAATGACTTCCAGACG.....ACTACTCACCCTACCG 1770

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 816138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued Patents NA: *
2: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
3: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
4: /cgn2_6/prodata/1/ina/6A_COMB.seq: *
5: /cgn2_6/prodata/1/ina/6B_COMB.seq: *
6: /cgn2_6/prodata/1/ina/6C_COMB.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1770	100.0	1770	3	US-08-191-160-3
2	1763.6	99.6	3750	3	US-08-191-160-22
3	1553.8	87.8	8638	4	US-10-029-907-7
4	1553.8	87.8	8638	4	US-10-029-907-25
5	1553.4	87.7	8642	4	US-10-029-907-2
6	1552.2	87.7	7989	4	US-09-539-601-10
7	1552.2	87.7	8001	4	US-09-539-601-7
8	1552.2	87.7	8637	4	US-09-539-601-4
9	1552.2	87.7	8638	4	US-10-029-907-6
10	1552.2	87.7	8638	4	US-10-029-907-24
11	1552.2	87.7	8639	4	US-10-029-907-1
12	1552.2	87.7	8643	4	US-10-029-907-4
13	1552.2	87.7	8649	4	US-09-539-601-13
14	1552.2	87.7	11076	4	US-09-539-601-1
15	1550.6	87.6	8648	4	US-10-029-907-5
16	1549	87.5	8001	4	US-09-539-601-22
17	1549	87.5	11076	4	US-09-539-601-25
18	1547.4	87.4	8001	4	US-09-539-601-16
19	1547.4	87.4	11076	4	US-09-539-601-19
20	1545.8	87.3	8001	4	US-09-539-601-28
21	1545.8	87.3	11076	4	US-09-539-601-31
22	1532.2	86.6	2991	1	US-08-324-977-39
23	1532.2	86.6	2991	2	US-08-384-616-49
24	1532.2	86.6	2991	3	US-08-904-686A-49
25	1532.2	86.6	2991	3	US-09-315-850-49
26	1532.2	86.6	7863	1	US-08-324-977-35
27	1532.2	86.6	7863	2	US-08-384-616-35

28	1532.2	86.6	7863	2	US-08-904-686A-35	Sequence 35, Appl
29	1532.2	86.6	7863	3	US-09-315-850-35	Sequence 35, Appl
30	1532.2	86.6	7917	1	US-08-324-977-31	Sequence 31, Appl
31	1532.2	86.6	7917	2	US-08-384-616-31	Sequence 31, Appl
32	1532.2	86.6	7917	2	US-08-904-686A-31	Sequence 31, Appl
33	1532.2	86.6	7917	3	US-09-315-850-31	Sequence 31, Appl
34	1532.2	86.6	9030	1	US-08-324-977-13	Sequence 13, Appl
35	1532.2	86.6	9030	2	US-08-384-616-13	Sequence 13, Appl
36	1532.2	86.6	9030	2	US-08-904-686A-13	Sequence 13, Appl
37	1532.2	86.6	9030	3	US-09-315-850-13	Sequence 13, Appl
38	1532.2	86.6	9416	1	US-08-324-977-1	Sequence 1, Appl
39	1532.2	86.6	9416	2	US-08-384-616-1	Sequence 1, Appl
40	1532.2	86.6	9416	2	US-08-904-686A-1	Sequence 1, Appl
41	1532.2	86.6	9416	3	US-09-315-850-1	Sequence 1, Appl
42	1532.2	86.6	9416	3	US-08-823-895A-27	Sequence 27, Appl
43	1529	86.4	9595	3	US-09-014-416-4	Sequence 4, Appl
44	1522.6	86.0	9472	4	US-08-150-204E-96	Sequence 96, Appl
45	1508.2	85.2	9413	4	US-09-827-688-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-08-191-160-3
; Sequence 3, Application US/08191160
; Patent No. 6210675
; GENERAL INFORMATION:
; APPLICANT: Highfield, Peter Edmund
; APPLICANT: Rodgers, Brian Colin
; APPLICANT: Tedder, Richard Seton
; APPLICANT: Barbara, John Anthony James
; TITLE OF INVENTION: Viral Agent
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESSES:
; ADDRESSER: Rothwell, Figg, Ernst & Kurz
; STREET: 1700 K Street
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy diskette, 5.25 inch, 360 Kb storage
; COMPUTER: IBM AT compatible
; OPERATING SYSTEM: MS-DOS V3.2
; SOFTWARE: Wordperfect 5.0 (DOS text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/191,160
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/628,516
; FILING DATE: 17 DEC 1990
; APPLICATION NUMBER: UK 89 28 562.1
; FILING DATE: 18 DEC 1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 90 04 414.0
; FILING DATE: 27 FEB 1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 90 04 814.1
; FILING DATE: 03 MAR 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: E. Anthony Figg
; REGISTRATION NUMBER: 27,195
; REFERENCE/DOCKET NUMBER: 1645-103A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 833-5740
; TELEFAX: (202) 833-5740
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1770 base pairs
; TYPE: nucleotide with corresponding protein
; STRANDEDNESS: single

```

; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to genomic RNA
; ORIGINAL SOURCE:
; ORGANISM: human; serum infectious for PT-NANBH
; IMMEDIATE SOURCE:
; LIBRARY: clone Jg2 from cDNA library in lambda gtl1
; FEATURE:
; LOCATION: from 1 to 1770 bp portion of the PT-NANBH
; LOCATION: polyprotein
; OTHER INFORMATION: probably encodes viral non-structural
; OTHER INFORMATION: proteins
US-08-191-160-3

```

```

Query Match      100.0%; Score 1770; DB 3; Length 1770;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 CMAATGACTTCCAGACGCTGACCTCATGAGGCCAACCCTCTGTGGCGCATGAGATG 60
DB 1 CMAAATGACTTCCAGACGCTGACCTCATGAGGCCAACCCTCTGTGGCGCATGAGATG 60
QY 61 GCGGGGACATTAACCGCGGTGAGTGCAGAAACAAGTAGTAATCCTGTGGCGCATGAGATG 120
DB 61 GCGGGGACATTAACCGCGGTGAGTGCAGAAACAAGTAGTAATCCTGTGGCGCATGAGATG 120
QY 121 CGGCTCCGAGCGAGAGAGATGAGCGGGAAGTGTCCGTCGCGCGAGATCCTGGGAAA 180
DB 121 CGGCTCCGAGCGAGAGAGATGAGCGGGAAGTGTCCGTCGCGCGAGATCCTGGGAAA 180
QY 181 TCCAGAAATTTCCACACAGCGATGCCCGCATGGGCAACCGCGAGATTACAACCTTCGCTG 240
DB 181 TCCAGAAATTTCCACACAGCGATGCCCGCATGGGCAACCGCGAGATTACAACCTTCGCTG 240
QY 241 CTGAGTCTCTGAGAGGCGCGGACTACGTCCTCCAGTGTATCATGGGTGCCCATGGCA 300
DB 241 CTGAGTCTCTGAGAGGCGCGGACTACGTCCTCCAGTGTATCATGGGTGCCCATGGCA 300
QY 301 CCTACTAAGACCCCTCTATATACACCTCCACGAGAAAGAGACAGTTGTTCTGACAGA 360
DB 301 CCTACTAAGACCCCTCTATATACACCTCCACGAGAAAGAGACAGTTGTTCTGACAGA 360
QY 361 TCCACCGTGTCTTTCGCTGCGCGAGAGCTTCCCAAAAGGCTTTTGTGATGCTCCGACCG 420
DB 361 TCCACCGTGTCTTTCGCTGCGCGAGAGCTTCCCAAAAGGCTTTTGTGATGCTCCGACCG 420
QY 421 TCGGCGGTGACAGCGGCAACCGGCAACCGGCTCCGACCAATTCCTCCGACGACGCGGA 480
DB 421 TCGGCGGTGACAGCGGCAACCGGCAACCGGCTCCGACCAATTCCTCCGACGACGCGGA 480
QY 481 GCGAGATCTGAGCGTGTGATTCCTTCATGCCCCCTTGTAGGGGAGCCGGGGAGC 540
DB 481 GCGAGATCTGAGCGTGTGATTCCTTCATGCCCCCTTGTAGGGGAGCCGGGGAGC 540
QY 541 CCCGATCTGAGCGAGGCTTGTGTCTACCGGTGAGAGAGCCGGTGAAGACGTCGTC 600
DB 541 CCCGATCTGAGCGAGGCTTGTGTCTACCGGTGAGAGAGCCGGTGAAGACGTCGTC 600
QY 601 TGCCTCTGATGCTCTACATGAGAGAGGCTCTGATACCGCAATGCGTCGAGAGGA 660
DB 601 TGCCTCTGATGCTCTACATGAGAGAGGCTCTGATACCGCAATGCGTCGAGAGGA 660
QY 661 AGCAAGCTGCCCATCAACGCGTTGAGCAACTTTTGTGCGTCACACACAATGATCTAC 720
DB 661 AGCAAGCTGCCCATCAACGCGTTGAGCAACTTTTGTGCGTCACACACAATGATCTAC 720
QY 721 GCTACCAATCCCGAGCGCAAGCCAGCGGCAAGAGGTACACTTTTGAAGACTGCA 780
DB 721 GCTACCAATCCCGAGCGCAAGCCAGCGGCAAGAGGTACACTTTTGAAGACTGCA 780
QY 781 ATCCGTGACGATCATCAACGAGACGTCGCAAGAGATGAAGGCGAAGCGTCACAGTT 840
DB 781 ATCCGTGACGATCATCAACGAGACGTCGCAAGAGATGAAGGCGAAGCGTCACAGTT 840

```

```

QY 841 AAGGCTAAGCTTCTATCATGAGAGAGAACCTGCAAGCTGAGCGCCCACTTTCGGCCAAA 900
DB 841 AAGGCTAAGCTTCTATCATGAGAGAGAACCTGCAAGCTGAGCGCCCACTTTCGGCCAAA 900
QY 901 TCTAAATTTGGCTATGAGGCGCAAAAGAGCTCCGGAACCTATCTCAGCAAGGCCATTTAACAC 960
DB 901 TCTAAATTTGGCTATGAGGCGCAAAAGAGCTCCGGAACCTATCTCAGCAAGGCCATTTAACAC 960
QY 961 ATCCGCTCCGTGTGGAGAGACTTGTGTGAGACACTGTAACACCAATTTGACACCACTC 1020
DB 961 ATCCGCTCCGTGTGGAGAGACTTGTGTGAGACACTGTAACACCAATTTGACACCACTC 1020
QY 1021 ATGCAAAAAATGAGGTTTTCGCTCCACCAAGAGAGAGGCGCCGACAGCTTCG 1080
DB 1021 ATGCAAAAAATGAGGTTTTCGCTCCACCAAGAGAGAGGCGCCGACAGCTTCG 1080
QY 1081 CTATTCGTTTCCAGACTTGGGGGTGCGTGTGTGCGAGAAATGGCCCTCTATGACGTG 1140
DB 1081 CTATTCGTTTCCAGACTTGGGGGTGCGTGTGTGCGAGAAATGGCCCTCTATGACGTG 1140
QY 1141 GTCTCACCCCTCCCTCAGGCTGTGATGGGCTCCTCGTACGGAATTCAGTATTTCTCTGGA 1200
DB 1141 GTCTCACCCCTCCCTCAGGCTGTGATGGGCTCCTCGTACGGAATTCAGTATTTCTCTGGA 1200
QY 1201 CAGCGGCTGAGATTCTGTGTGAACGCTGGAATCAAGAAAGACCCCTATGGGCTTTGCA 1260
DB 1201 CAGCGGCTGAGATTCTGTGTGAACGCTGGAATCAAGAAAGACCCCTATGGGCTTTGCA 1260
QY 1261 TATGACACCCGCTGTTTGAATCAACAGTCACTGAGAAATGACATCCGTTGAGAGATCA 1320
DB 1261 TATGACACCCGCTGTTTGAATCAACAGTCACTGAGAAATGACATCCGTTGAGAGATCA 1320
QY 1321 ATTATCAATGTGTGACTTGTGAGGCGCCGAGACGAGACGAGCATTAAGTCCGTCACAGAG 1380
DB 1321 ATTATCAATGTGTGACTTGTGAGGCGCCGAGACGAGACGAGCATTAAGTCCGTCACAGAG 1380
QY 1381 CGGCTTTATATCGGAGGCTCCCTGATTAATCAAAAGGCGCAAACTGGCGCTATCGCGG 1440
DB 1381 CGGCTTTATATCGGAGGCTCCCTGATTAATCAAAAGGCGCAAACTGGCGCTATCGCGG 1440
QY 1441 TCCCGCGGAGCGGCTGTCTGACGATAGCTGCGGTATACCTTCACATGTTACTTGAAG 1500
DB 1441 TCCCGCGGAGCGGCTGTCTGACGATAGCTGCGGTATACCTTCACATGTTACTTGAAG 1500
QY 1501 GCTCTGACGCTGTGACGCTGCAAGAGCTCCAGGACTGCAAGATGCTGTGTGCGAGAGAC 1560
DB 1501 GCTCTGACGCTGTGACGCTGCAAGAGCTCCAGGACTGCAAGATGCTGTGTGCGAGAGAC 1560
QY 1561 GGCCTTGTCTTATCTGTGAGAGCGCGGGAACCCAGAGAGAGCGCGAGGCTACGAGTC 1620
DB 1561 GGCCTTGTCTTATCTGTGAGAGCGCGGGAACCCAGAGAGAGCGCGAGGCTACGAGTC 1620
QY 1621 TTCAACGAGGCTATGACTAGTACTTGTGCCCCCGGGAGCCGCCCCCAACGAAATAC 1680
DB 1621 TTCAACGAGGCTATGACTAGTACTTGTGCCCCCGGGAGCCGCCCCCAACGAAATAC 1680
QY 1681 GACCTGAGGTGATTAACATGCTCCCTCAATGCTGCGGCGGCAAGATGCACTTGGC 1740
DB 1681 GACCTGAGGTGATTAACATGCTCCCTCAATGCTGCGGCGGCAAGATGCACTTGGC 1740
QY 1741 AAAAGGATATCACTCACTCACCCGTCACCG 1770
DB 1741 AAAAGGATATCACTCACTCACCCGTCACCG 1770

```

```

RESULT 2
US-08-191-160-22
; Sequence 22, Application US/08191160
; Patent No. 6210675
; GENERAL INFORMATION:
; APPLICANT: Highfield, Peter Edmund
; APPLICANT: Rodgers, Brian Colin
; APPLICANT: Tedder, Richard Seton

```

APPLICANT: Barbara, John Anthony James
 TITLE OF INVENTION: Viral Agent
 NUMBER OF SEQUENCES: 25
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Rothwell, Figg, Ernst & Kurz
 STREET: 1700 K Street
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20006
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy diskette, 5.25 inch, 360 Kb storage
 COMPUTER: IBM AT compatible
 OPERATING SYSTEM: MS-DOS V3.2
 SOFTWARE: Wordperfect 5.0 (DOS text)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/191,160
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/628,516
 FILING DATE: 17 DEC 1990
 APPLICATION NUMBER: UK 89 28 562.1
 FILING DATE: 18 DEC 1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: UK 90 04 414.0
 FILING DATE: 27 FEB 1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: UK 90 04 814.1
 FILING DATE: 03 MAR 1990
 ATTORNEY/AGENT INFORMATION:
 NAME: E. Anthony Figg
 REGISTRATION NUMBER: 27,195
 REFERENCE/DOCKET NUMBER: 1645-103A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 833-5740
 TELEFAX: (202) 833-5744
 INFORMATION FOR SEQ ID NO: 22:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3750 base pairs
 TYPE: nucleotide with corresponding protein
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to genomic RNA
 ORIGINAL SOURCE:
 ORGANISM: human; serum infectious for PT-NANBH
 IMMEDIATE SOURCE:
 LIBRARY: cDNA clones from 3' end of the genome
 FEATURE:
 LOCATION: from 1 to 3750 bp portion of the PT-NANBH
 LOCATION: polyprotein
 OTHER INFORMATION: viral non-structural proteins
 US-08-191-160-22
 Query Match 99.8%; Score 1763.6; DB 3; Length 3750;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1766; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 CAAAATGACTTCCAGACGCTGACCTCATGAGAGCCAACTCTGTGAGCGGATGAGATG 60
 DB 1981 CAAAATGACTTCCAGACGCTGACCTCATGAGAGCCAACTCTGTGAGCGGATGAGATG 2040
 QY 61 GGGGGGAGCATTAACCGCGGTGAGTCAAGAGAACAGGTAGTAATCTTGAATCTTTTGAC 120
 DB 2041 GGGGGGAGCATTAACCGCGGTGAGTCAAGAGAACAGGTAGTAATCTTGAATCTTTTGAC 2100
 QY 121 CCGCTCCGAGCGAGAGATGAGCGGAGAGTCTCCGCGGAGAGATCCGCGGAAA 180
 DB 2101 CCGCTCCGAGCGAGAGATGAGCGGAGAGTCTCCGCGGAGAGATCCGCGGAAA 2160
 QY 181 TCCAGAAATTTCCACAGCGATGCCGATGGGCAAGCGCGGATTAACAACCTCCGCTG 240
 DB 2161 TCCAGAAATTTCCACAGCGATGCCGATGGGCAAGCGCGGATTAACAACCTCCGCTG 2220

241 CTGAGATCTGGAAAGGCCCCGAGACTACGTCCTCCAGTGTATCATGGGTGCCACTGCCA 300
 DB 2221 CTGAGATCTGGAAAGGCCCCGAGACTACGTCCTCCAGTGTATCATGGGTGCCACTGCCA 2280
 QY 301 CCTACTAAGACCCCTCTCTATTAACAACCTCCAGAGAGAGAGACAGTTGTTGACAGAA 360
 DB 2281 CCTACTAAGACCCCTCTCTATTAACAACCTCCAGAGAGAGAGACAGTTGTTGACAGAA 2340
 QY 361 TCCACCGTGTCTTTCGCTCCGAGAGTTCGCCAACAAGCTTTTGTAGTCTCCGACCG 420
 DB 2341 TCCACCGTGTCTTTCGCTCCGAGAGTTCGCCAACAAGCTTTTGTAGTCTCCGACCG 2400
 QY 421 TCGGCGGTGAGACAGGCGGACCGGCAACCGGCGCTCCGACCAATTCCTCCAGACGCGG 480
 DB 2401 TCGGCGGTGAGACAGGCGGACCGGCAACCGGCGCTCCGACCAATTCCTCCAGACGCGG 2460
 QY 481 GCAGATCTGAGGTTGAGTGTATTCCTCATATGCCCCCTTTGAGGGGAGCCGGGGGAC 540
 DB 2461 GCAGATCTGAGGTTGAGTGTATTCCTCATATGCCCCCTTTGAGGGGAGCCGGGGGAC 2520
 QY 541 CCGATCTCAGCGACGGGTCTTGTCTACCGTGAAGTGAAGGCGGCTGAGACGTCTGTC 600
 DB 2521 CCGATCTCAGCGACGGGTCTTGTCTACCGTGAAGTGAAGGCGGCTGAGACGTCTGTC 2580
 QY 601 TGTGCTGATGTCTCTACACATGAGACAGGCGCTGTGATACGCGCATGCGCTGGAGGAA 660
 DB 2581 TGTGCTGATGTCTCTACACATGAGACAGGCGCTGTGATACGCGCATGCGCTGGAGGAA 2640
 QY 661 AGCAAGCTCCCATCAACGCGTGTGAGCACTCTTGTCTGTCACCAACATGCTCTAC 720
 DB 2641 AGCAAGCTCCCATCAACGCGTGTGAGCACTCTTGTCTGTCACCAACATGCTCTAC 2700
 QY 721 GCTACCAATCCCGGACGCGCAAGCGGACAGGAGAGTCACTTTGACAGATGCGAA 780
 DB 2701 GCTACCAATCCCGGACGCGCAAGCGGACAGGAGAGTCACTTTGACAGATGCGAA 2760
 QY 781 ATCCGTGAGATCACTACAGAGACGCTCAAGGAGATGAGGCGCAAGGCGTCCACAGTT 840
 DB 2761 ATCCGTGAGATCACTACAGAGACGCTCAAGGAGATGAGGCGCAAGGCGTCCACAGTT 2820
 QY 841 AAGGCTAAGCTTCTATCACTAGAGAGAGCTGCAAGCTGACGCCCCACATTCGCGCAAA 900
 DB 2821 AAGGCTAAGCTTCTATCACTAGAGAGAGCTGCAAGCTGACGCCCCACATTCGCGCAAA 2880
 QY 901 TCTAAATTTGGCTATGAGGCAAGAGAGTCCGGAACCTATCCAGCAAGGCAATTAACAC 960
 DB 2881 TCTAAATTTGGCTATGAGGCAAGAGAGTCCGGAACCTATCCAGCAAGGCAATTAACAC 2940
 QY 961 ATCCGCTCCGTGTGAGAGAGCTTGTGAGAGACATGGAACCAATTTGACACCAACATC 1020
 DB 2941 ATCCGCTCCGTGTGAGAGAGCTTGTGAGAGACATGGAACCAATTTGACACCAACATC 3000
 QY 1021 ATGGCAAAATATGAGGTTTTTGTGCTCAACCAAGAGAGAGAGGCGCAAGCGAGCTCGC 1080
 DB 3001 ATGGCAAAATATGAGGTTTTTGTGCTCAACCAAGAGAGAGAGGCGCAAGCGAGCTCGC 3060
 QY 1081 CTATGCTGTTCACAGACTTGGGGGTCGCTGTGTGTGAGAAATATGAGGCTCTATGACGTG 1140
 DB 3061 CTATGCTGTTCACAGACTTGGGGGTCGCTGTGTGTGAGAAATATGAGGCTCTATGACGTG 3120
 QY 1141 GTTTCACCTCCCTCAAGCTGTGATGAGGCTCTCTGATGAGATTCAGATTTCTCTGGA 1200
 DB 3121 GTTTCACCTCCCTCAAGCTGTGATGAGGCTCTCTGATGAGATTCAGATTTCTCTGGA 3180
 QY 1201 CAGCGGATGAGATTTCTGTGTAAGCGCTGAAATCAAGAAGACCCCTATGAGCTTTGGA 1260
 DB 3181 CAGCGGATGAGATTTCTGTGTAAGCGCTGAAATCAAGAAGACCCCTATGAGCTTTGGA 3240
 QY 1261 TATGACACCGCTGTTTGAATCAACAGTCACTGAGATGACATCCGTGTAGAGAGTCA 1320
 DB 3241 TATGACACCGCTGTTTGAATCAACAGTCACTGAGATGACATCCGTGTAGAGAGTCA 3300

```

QY 1321 ATTATCAATGTTGACTGTGGCCCCCGAAGCCAGACAGTCAAGTCTGCTCAGAG 1380
|
|
|
Db 3301 ATTATCAATGTTGACTGTGGCCCCCGAAGCCAGACAGTCAAGTCTGCTCAGAG 3360
|
|
|
QY 1381 CGGCTTATATGCGGGGTCCCTGACTAATTCAAAAGGCGAGAACTGCGGCTATGCGCG 1440
|
|
|
Db 3361 CGGCTTATATGCGGGGTCCCTGACTAATTCAAAAGGCGAGAACTGCGGCTATGCGCG 3420
|
|
|
QY 1441 TCCCGCGAGCGCGCTGTGACGACTAGCTCGGTAATACCTCAATGTTACTTTGAAG 1500
|
|
|
Db 3421 TCCCGCGAGCGCGCTGTGACGACTAGCTCGGTAATACCTCAATGTTACTTTGAAG 3480
|
|
|
QY 1501 GCCTTGACGCTGTGACCTGCAAGAGCTCCAGAGCTGACGATGCTGTGTGCGAGAC 1560
|
|
|
Db 3481 GCCTTGACGCTGTGACCTGCAAGAGCTCCAGAGCTGACGATGCTGTGTGCGAGAC 3540
|
|
|
QY 1561 GGGCTTGTGCTTATCTGTGAGAGCGCGGGAACCCAGAGAGAGCGCGGAGCTTACGAGTC 1620
|
|
|
Db 3541 GGGCTTGTGCTTATCTGTGAGAGCGCGGGAACCCAGAGAGAGCGCGGAGCTTACGAGTC 3600
|
|
|
QY 1621 TTCAAGGAGGCTATGACTAGTACTGTGCCCCCGGAGACCCGCGCCAAACAGAAATAC 1680
|
|
|
Db 3601 TTCAAGGAGGCTATGACTAGTACTGTGCCCCCGGAGACCCGCGCCAAACAGAAATAC 3660
|
|
|
QY 1681 GACTGAGAGTATTAACATCATGCTCTTCCAAATGTGTGCTGCGCAGCATGCTGTGC 1740
|
|
|
Db 3661 GACTGAGAGTATTAACATCATGCTCTTCCAAATGTGTGCTGCGCAGCATGCTGTGC 3720
|
|
|
QY 1741 AAAAGGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1770
|
|
|
Db 3721 AAAAGGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3750
|
|
|

```

```

RESULT 3
US-10-029-907-7
; Sequence 7, Application US/10029907
; Patent No. 6706874
; GENERAL INFORMATION:
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
; TITLE OF INVENTION: HEPATITIS C VIRUS
; FILE REFERENCE: 13/083
; CURRENT APPLICATION NUMBER: US/10/029,907
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,857
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PaeCSO for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 8638
; TYPE: DNA
; ORGANISM: HCV
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1802)...(8407)
US-10-029-907-7

```

```

Query Match      87.8%; Score 1553.8; DB 4; Length 8638;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1633; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

```

```

QY 5 ATGACTTCCAGACGCTGAACCTCATGAGCCAACTCTGTGGCGGAGATGAGCG 64
|
|
|
Db 6030 ATGACTCCCCGACGCTGACCTCATGAGCCAACTCTGTGGCGGAGATGAGCG 6089
|
|
|
QY 65 GGGACATTAACCCGCTGTGAGTCAAGAGACAAAGTATTAATCTTGAATCTTTGACCGC 124
|
|
|
Db 6090 GGAACATCAACCCGCTGTGAGTCAAGAGATTAAGTATTAATTTTGAATCTTTGACCGC 6149
|
|
|
QY 125 TCCGAGCGGAGAGATGAGCGGGAAGTCCGTCGCCGCGAGATCCGCGGAATCCA 184
|
|
|
Db 6150 TCCAGCGGAGAGATGAGCGGGAAGTCCGTCGCCGCGAGATCCGCGAGAGTCCA 6209
|
|
|

```

```

QY 185 AGAAATTCACACAGCGATGCGCGATGAGGACCGCCGGAATTAACAACCTCCGCTGCTG 244
|
|
|
Db 6210 GGAATTCCTCGAGCGATGCGCGATGAGGACCGCCGGAATTAACAACCTCCGCTGCTG 6269
|
|
|
QY 245 AGTCTGGAAGAGCCCGGACTTACGTCCTTCCAGTGTATCATGGGTGCCCACTGCCACTA 304
|
|
|
Db 6270 AGTCTGGAAGAGCCCGGACTTACGTCCTTCCAGTGTATCATGGGTGCCCACTGCCACTG 6329
|
|
|
QY 305 CTAAGACCCCTCTTATACACCTCCACGGAAGAGAGACAGTGTCTTGAAGAAATCCA 364
|
|
|
Db 6330 CTAAGACCCCTCTTATACACCTCCACGGAAGAGAGACAGTGTCTTGAAGAAATCCA 6389
|
|
|
QY 365 CCGTGTCTTGTGCCCTGCGAGAGCTTGCACAAAAGCTTTTGTAGTCTTCGAGCCGTGCG 424
|
|
|
Db 6390 CCGTGTCTTGTGCCCTTGTGCGAGAGCTTCGACAAAAGACTTTCGCGAGCTCCGAATGTGCG 6449
|
|
|
QY 425 CCGTGAAGAGGCGACCGGCAACCGCCCTCTTGAACCAATCTTCGAGCGAGCGGAGAG 484
|
|
|
Db 6450 CCGTGAAGAGGCGACCGGCAACCGCCCTCTTGAACCAATCTTCGAGCGAGCGGAGAG 6509
|
|
|
QY 485 GATCTGAGGCTTGAAGTGTATTCCTCAATGCCCCCTTGAAGGAGAGCGCGGAGACCCCG 544
|
|
|
Db 6510 GATCTGAGGCTTGAAGTGTATTCCTCAATGCCCCCTTGAAGGAGAGCGCGGAGATCCCG 6569
|
|
|
QY 545 ATCTGAGGAGCGGCTTGTGTCTACCGTGAAGAGAGCGGTGAGAGCGTCTGTGCT 604
|
|
|
Db 6570 ATCTGAGGAGCGGCTTGTGTCTACCGTGAAGAGAGCGGTGAGAGCGTCTGTGCT 6629
|
|
|
QY 605 GCTGATGTCTTACATATGACAGGCGCTGTGATCAAGCCATGCGCTGCGAGAGAGCA 664
|
|
|
Db 6630 GCTGATGTCTTACATATGACAGGCGCTGTGATCAAGCCATGCGCTGCGAGAGAGCA 6689
|
|
|
QY 665 AGCTGCCCATCAAGCGGCTTGAAGCAACTCTTGTGCGGCTGACCAACAATGATTAAGCTTA 724
|
|
|
Db 6690 AGCTGCCCATCAAGCGGCTTGAAGCAACTCTTGTGCGGCTGACCAACAATGATTAAGCTTA 724
|
|
|
QY 725 CCACATCCCGGAGCGCAAGCGGCGAGAAAGTCACTTGTGACAGACTGCAATCC 784
|
|
|
Db 6750 CCACATCCCGGAGCGCAAGCGGCGAGAAAGTCACTTGTGACAGACTGCAATCC 784
|
|
|
QY 785 TGAAGATCACTTACAGGAGCTGTCTCAAGAGATGAAGGCGAAGCGTCCACAGTTAAG 844
|
|
|
Db 6810 TGAAGATCACTTACAGGAGCTGTCTCAAGAGATGAAGGCGAAGCGTCCACAGTTAAG 844
|
|
|
QY 845 CTAAGCTTCTATGAGTGAAGAGCGCTGCAAGCTGACGCGCCCACTTGGCGCAATCTA 904
|
|
|
Db 6870 CTAAGCTTCTATGAGTGAAGAGCGCTGCAAGCTGACGCGCCCACTTGGCGCAATCTA 904
|
|
|
QY 905 AATTTGGCTATGAGGCAAAAGAGCGTCCGGAACCTATCCAGAGGCAATTAACCAATCC 964
|
|
|
Db 6930 AATTTGGCTATGAGGCAAAAGAGCGTCCGGAACCTATCCAGAGGCAATTAACCAATCC 964
|
|
|
QY 965 GCTCCGTGTGAGAGAGCTTGTGGAAGACATGAAACCAATTGACCAACCATCATG 1024
|
|
|
Db 6990 GCTCCGTGTGAGAGAGCTTGTGGAAGACATGAAACCAATTGACCAACCATCATG 1024
|
|
|
QY 1025 CAAAAAATGAGATTTTGTGCGTCCAAACAGAGAGAGCGGCAAGCACTGCGCTTA 1084
|
|
|
Db 7050 CAAAAAATGAGATTTTGTGCGTCCAAACAGAGAGAGCGGCAAGCACTGCGCTTA 1084
|
|
|
QY 1085 TCGTGTCCAGACTTGGGAGGCTCGTGTGCGAGAAATGAGCCCTTATGACGTGTCT 1144
|
|
|
Db 7110 TCGTGTCCAGACTTGGGAGGCTCGTGTGCGAGAAATGAGCCCTTATGACGTGTCT 1144
|
|
|
QY 1145 CCACCTTCCCTCAGGCTGTATGAGGCTCTGTACGATTCAGATTTCTCTGACAGC 1204
|
|
|
Db 7170 CCACCTTCCCTCAGGCTGTATGAGGCTCTGTACGATTCAGATTTCTCTGACAGC 1204
|
|
|
QY 1205 GGGTGAAGTTCCTGTGTAAGCGCTGGAATCAAGAGAGACCCCTATGAGGCTTTGCAATG 1264
|
|
|
Db 7230 GGGTGAAGTTCCTGTGTAAGCGCTGGAATCAAGAGAGACCCCTATGAGGCTTTGCAATG 1264
|
|
|
QY 1265 AACCCGCTGTTTGAATCAACAGTCACTGAATGACATCCGTGTAGAGAGTCAATTT 1324
|
|
|

```

Db 7290 ACACCCGCTGTTTGACTCAACGCTCACTGAGATACATCCCTGTGAGAGATCATATCT 7349
Qy 1325 ATCAATGTTGTGACTTGGCCCCCGAAGCAGACAGGCGCATTAAGTGTCTTCAAGAGCGGC 1384
Db 7350 ACCAATGTTGTGACTTGGCCCCCGAAGCAGACAGGCGCATTAAGTGTCTTCAAGAGCGGC 7409
Qy 1385 TTTATATCGGGGGTCCCTGACTTAATTCAAAAGGGAGAACTGGCGCTATCGCGGTGGC 1444
Db 7410 TTTACATCGGGGGCCCCCTGACTTAATTCAAAAGGGAGAACTGGCGCTATCGCGGTGGC 7469
Qy 1445 GCGCAGCGCGCTGTGCTGACGACTAGTGGGTATATACCTTACATTTTAAAGGCT 1504
Db 7470 GCGCAGCGCGTGTGCTGACGACGACGCTGGGTATATACCTTACATTTTAAAGGCGC 7529
Qy 1505 CTGCAAGCTGTGAGCTGCAAGAGCTCAAGAGCTGCAAGTGTGTGTGCGAGAGCGGC 1564
Db 7530 CTGCGGCTGTGAGCTGCAAGAGCTCAAGAGCTGCAAGTGTGTGTGCGAGAGAGC 7589
Qy 1565 TTGTGTTTCTGTGAGAGCGGGAAACCAAGAGAGCGGGAGCTTACGATCTTCA 1624
Db 7590 TTGTGTTTCTGTGAGAGCGGGAAACCAAGAGAGCGGGAGCTTACGAGCTTCA 7649
Qy 1625 CCGAGGCTATGACTAGTACTGTGCCCCCGGGAGACCGCCCAACAGAAATACGAC 1684
Db 7650 CCGAGGCTATGACTAGTACTGTGCCCCCGGGAGACCGCCCAACAGAAATACGACT 7709
Qy 1685 TGGAGTTGATTAACATCATGCTCTCAATGTGTGTGTGCGGACGATGTGCGAAAA 1744
Db 7710 TGGAGTTGATTAACATCATGCTCTCAATGTGTGTGTGCGGACGATGTGCGAAAA 7769
Qy 1745 GGGTATCTAATCTACCCGCTGACCC 1769
Db 7770 GGGTATCTAATCTACCCGCTGACCC 7794

RESULT 4
US-10-029-907-25
Sequence 25, Application US/10029907
Patent No. 6706874
GENERAL INFORMATION:
APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
HEPATITIS C VIRUS
FILE REFERENCE: 13/083
CURRENT APPLICATION NUMBER: US/10/029,907
PRIOR FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 60/257,857
PRIOR FILING DATE: 2000-12-22
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 25
LENGTH: 8638
TYPE: DNA
ORGANISM: HCV
FEATURE:
NAME/KEY: CDS
LOCATION: (1802) ... (8407)
US-10-029-907-25

Query Match 87.8%; Score 1553.8; DB 4; Length 8638;
Best Local Similarity 92.5%; Pired. No. 0;
Matches 1633; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

Qy 5 ATGACTTCCAGACGCTGACCTCATGAGGCAACCTCTGTGGCGGACGAGATGAGCG 64
Db 6030 ATGACTTCCAGACGCTGACCTCATGAGGCAACCTCTGTGGCGGAGATGAGCG 6089
Qy 65 GGGACATTAACCGCGTGGAGTCAAGAAACAAGTAACTCTTGAAGCTCTTTCAGCCGC 124
Db 6090 GGAACATTAACCGCGTGGAGTCAAGAAACAAGTAACTCTTTCAGCCGC 6149
Qy 125 TCCGACCGAGAGAGATGAGCGGGAAGTGTCCCTCCGCGGAGATCTTCGGAATCCA 184

Db 6150 TCCAGCGGAGAGATGAGAGGAAGTATCCGTTCCGCGGAGATCTTCGGAAGTCCA 6209
Qy 185 AGAATTTCCACGAGGATGCCCGATGGGACAGCCCGGATTAACAACCTTCGCTGTG 244
Db 6210 GGAATTTCCCGTGAAGGATGCCCATATGGGACAGCCCGGATTAACAACCTTCGCTGTG 6269
Qy 245 AGTCCGGAAGGCCCCCGGATAGTGTCCCTCAAGTGTATAGGGTCCGACGCACTTA 304
Db 6270 AGTCCGGAAGGCCCCCGGATAGTGTCCCTCAAGTGTATAGGGTGTCAATGGCCGCTG 6329
Qy 305 CTAAAGCCCTCTCTATACCACTTCAAGAGAAAGAGACAGTGTCTTCAAGAAATCCA 364
Db 6330 CCAAGGCCCTCTCTATACCACTTCAAGAGAAAGAGACAGTGTCTTCAAGAAATCCA 6389
Qy 365 CCGTGTCTTCTGCTGCGGAGCTTGTGCAAAAGGCTTTGTGTAGTCTCGACGCTCG 424
Db 6390 CCGTGTCTTCTGCTGCGGAGCTTGTGCAAAAGGCTTTGTGTAGTCTCGAATCGTCCG 6449
Qy 425 CCGTGTGAGCGGACGCGAACCGCCCTCTGACCAATCTTCCGACGAGCGAGCAG 484
Db 6450 CCGTGTGAGCGGACGCGAACCGCCCTCTGACCAATCTTCCGACGAGCGAGCAGC 6509
Qy 485 GATCTGACGTTGATGTATTTCTCCATGCCCCCTTGAAGGAGGAGCGGAGACCCG 544
Db 6510 GATCTGACGTTGATGTATTTCTCCATGCCCCCTTGAAGGAGGAGCGGAGACCCG 6569
Qy 545 ATCTCAGCAGCGGCTTGTGTCTTACCGTGAAGAGGCGGTGAGAGCTGTCTGCT 604
Db 6570 ATCTCAGCAGCGGCTTGTGTCTTACCGTGAAGAGGCGGTGAGAGCTGTCTGCT 6629
Qy 605 GCTTCATGTCTTACATATGAGAGGCTCTGATACCGCATGCGGTGGAGAGAAAGA 664
Db 6630 GCTTCATGTCTTACATATGAGAGGCTCTGATACCGCATGCGGTGGAGAGAAAGA 6689
Qy 665 AGCTGCCATCAACGCGTTGAGCAACTCTTGTGCTGCTACCAACAATGTCTACGCTA 724
Db 6690 AGCTGCCATCAATGAGCTGAGCAACTCTTGTGCTGCTACCAACAATGTGTATGCTA 6749
Qy 725 CCAACATCCGCGAGCGCAACGCGGAGAGAAAGTCACTTTGACAGACTTGCATAATCC 784
Db 6750 CCAACATCCGCGAGCGCAACGCGGAGAGAAAGTCACTTTGACAGACTTGCAGTCC 6809
Qy 785 TGGAGATCACTTACAGAGCGTCTCAAGAGATGAAGGCGAAAGGTCTCAAGTTAAG 844
Db 6810 TGGAGATCACTTACAGAGCGTCTCAAGAGATGAAGGCGAAAGGTCTCAAGTTAAG 849
Qy 845 CTAAGCTTCAATCAAGTGAAGAGCGTCAAGCTGACGCGCCCACTTGGCCAAATCTA 904
Db 6870 CTAAGCTTCAATCAAGTGAAGAGCGTCAAGCTGACGCGCCCACTTGGCCAAATCTA 6929
Qy 905 AATTGGCTATGAGGCAAGAGAGTCCGGAACCTATCCAGAAAGGCAATTAACACATCC 964
Db 6930 AATTGGCTATGAGGCAAGAGAGTCCGGAACCTATCCAGAAAGGCGTTAACACATCC 6989
Qy 965 GCTCCGTGTGGAGAGCTTGTGGAAGACACTGAAACACCAATTTGACACCACTATAG 1024
Db 6990 GCTCCGTGTGGAGAGCTTGTGGAAGACACTGGAAGACCAATTTGACACCACTATAG 7049
Qy 1025 CAAAAATGAGGTTTTCTGCTCAACAGAGAGAGAGGCGCAAGCAGCTGCGCTTA 1084
Db 7050 CAAAAATGAGGTTTTCTGCTCAACAGAGAGAGAGGCGCAAGCAGCTGCGCTTA 7109
Qy 1085 TCGTGTTCAGAGCTTGGGGGTCCTGTGTGCGAGAAATGAGCCCTTATGAGCGTGTCT 1144
Db 7110 TCGTGTTCAGAGCTTGGGGGTCCTGTGTGCGAGAAATGAGCCCTTATGAGTGTGTCT 7169
Qy 1145 CCAACCTCCCTCAGAGTGTGAGTGGGTCTCTGTAAGGATTTCAAGTATTTCTCGAGCAGC 1204
Db 7170 CCAACCTCCCTCAGAGGCGGTGAGTGGGTCTTTCATACGATTTCAAAATCTTCTCGAGCAGC 7229
Qy 1205 GGGTGAAGTCTCTGAGAGCGCTGGAATCAAAAGAGACCCCTATGAGGCTTTCATATG 1264

Db	7230	GGGTGCAATTCCTGCTGAATGCTGTGGAAGCGAAAGAAATGCCCTTATGGGCTTCGCAATATG	7289
OY	1265	ACACCCGCTGTTTGTGACTCAACAGTCACTGAGAAATGACATCCGTGTAGAGGATCAATT	13244
Db	7290	ACACCCGCTGTTTGTGACTCAACAGGTCATGGAATGATCATCCGTGTGAGGATCAATCT	7349
OY	1325	ATCATATGTTTGTGACTTTGGCCCCCGAAGCAGACAGGCCATTAAGTTCGCTCACAGCGGC	1384
Db	7350	ACCAATGTTTGTGACTTTGGCCCCCGAAGCAGACAGGCCATTAAGTTCGCTCACAGCGGC	7409
OY	1385	TTTATATGGGGGGTCCCTGACTAATTCAAAAAGGCAAGACTGGGGCTATGCGCGATGC	1444
Db	7410	TTTATATGGGGGGCCCCCTGACTAATTTTAAAGGCAAGACTGGGGCTATGCGCGATGC	7469
OY	1445	GCGCGAGCGCGGTGTGACGACTAGCTGCGGTAAATCCCTCACATGTTACTTGAAGGCT	1504
Db	7470	GCGCGAGCGGTGTGACTAGCAGACGAGCTGGGTAAATCCCTCACATGTTACTTGAAGGCG	7529
OY	1505	CTGCAAGCTGTTCGAGCTGCAAAAGCTCCAGAGCTGCAGATCTGTGTGCGGAGCGGCC	1564
Db	7530	CTGCGGGCTGTTCGAGCTGCGAAGCTCCAGAGCTGCAGATCTGTGTGCGGAGCGGCC	7589
OY	1565	TTTGTCGTTATCTGTGAGAGCGCGGGGAACCCAGAGAGACGCGGCGAGCTTAAGATCTTCA	1624
Db	7590	TTTGTCGTTATCTGTGAAAGCGCGGGAGCCCAAGAGAGAGAGGCGAGCTTAAGATCTTCA	7649
OY	1625	CGAGAGCTATAGTACTAGGACTCTGCGCCCCCCCCCGGGGAGCCGCGCCCAACAGAAATACGACC	1684
Db	7650	CGAGAGCTATAGTACTAGATATCTTGCCCCCCCCCTGGGGAGCCGCGCCCAACAGAAATACGACT	7709
OY	1685	TGAGATTGATTAACATCATGCTCTTCCAAATGTTGTGCGTGCAGACAGATGCATCTGGCAAA	1744
Db	7710	TGAGATTGATTAACATCATGCTCTTCCAAATGTTGTGCGTGCAGACAGATGCATCTGGCAAA	7769
OY	1745	GGGTATACTACTCAACCCGTGACCC	1769
Db	7770	GGGTATACTATCTCAACCCGTGACCC	7794

```

RESULT 5
US-10-029-907-2
Sequence 2, Application US/10029907
Patent No. 6706874
GENERAL INFORMATION:
APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
TITLE OF INVENTION: HEPATITIS C VIRUS
FILE REFERENCE: 13/083
CURRENT APPLICATION NUMBER: US/10/029,907
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 60/257,857
PRIOR FILING DATE: 2000-12-22
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 2
LENGTH: 8642
TYPE: DNA
ORGANISM: HCV
FEATURE:
NAME/KEY: CDS
LOCATION: (1802) ..(8407)
NAME/KEY: variation
LOCATION: 6268
OTHER INFORMATION: r = a or g
NAME/KEY: variation
LOCATION: 4446
OTHER INFORMATION: r = a or g
US-10-029-907-2

```

Query Match	87.8%;	Score 1553.4;	DB 4;	Length 8642;
Best Local Similarity	92.5%;	Pred. No. 0;		
Matches 1632;	Conservative	1;	Mismatches 132;	Indels 0;
			Gaps	0

QY	5	ATGATCTTCCAGACGCTGATCTCATCGAGGCCAACCCTCTGTGGCGGCACTGAGATGGGCG	64
Db	6030	ATGATCTCCCGGACCGCTGATCTCATCGAGGCCAACCTCTGTGGCGGCACTGAGATGGGCG	6083
QY	65	GGGACATTACCCGCGTGGAGTGCAGGAACAAGTGTATATCTGGACTCTTTTCGACCGCG	124
Db	6090	GGACATACCCGCGTGGAGTGCAGGAATTAAGTATGTATTTTGGACTCTTTTCGAGCCGC	6143
QY	125	TCCGAGCGGAGAGGATAGCGGGAGTGTCCGTCCCGGCGAGATTCCTGCGGAATCA	184
Db	6150	TCGAAGCGGAGAGGATAGGAGGAGTATCCGTTCCGCGGAGATTCCTGCGGAGTCCA	6203
QY	135	AGAAATTTCCACACAGCGATGCCCGCATGGGCAACGCCGGAATTACACCTCCGCTGCG	244
Db	6210	GGAAATTTCCCTCGAGCGATGCCCATATGGGACCGCCGGAATTACAACTCTCACTGTTTG	6265
QY	245	AGTCCTGGAAAGGCCCGGACTTACGCTCCCTCCAGTGTATCATGGGGCCCATGCGCCACTA	304
Db	6270	AGTCCTGGAAAGGACCCGGACTTACGCTCCCTCCAGTGTATCAAGGATGTCAATTGCGCCCTG	6322
QY	305	CTAAGACCCCTCTATATACCACTCCACGAGAAAGAGGACAGTTGTTCTGACAGATCCA	364
Db	6330	CCAAGGCCCTCCGATATACCACTCCACGAGAAAGAGGACGTTGTCTGTGTACAGATCTCA	6383
QY	365	CCGTGTCTTTCGCGCTGGCGGAGCTTGCCACAAAGCTTTTGTATGCTCCGACGCTGCG	424
Db	6390	CCGTGTCTTTCGCGCTGGCGGAGCTTGCCACAAAGACTTTCGGGCACTCCGAATGTGCG	6448
QY	425	CCGTGCAGAGGGGACGGGCAACCGCCCTCTGACCAATCTCTCCGACGACGGCGGAGGAG	484
Db	6450	CCGTGCAGAGGGGACGGGCAACCGCTCTCTGACCAACCTCTCCGACGACGAGCGGCG	6505
QY	485	GATCGACGTTGAGTGTATTCCTTCATAGCCCCCTTGAGGGGGAGCGGGGAGCCCG	544
Db	6510	GATCGACGTTGAGTGTATTCCTTCATAGCCCCCTTGAGGGGGAGCGGGGAGATCCCG	6565
QY	545	ATCTCAGGACGGGTCTTGTGTCTACCGTAGTAGAGGACCGGTGAGGACGTCTTGCT	604
Db	6570	ATCTCAGGACGGGTCTTGTGTCTACCGTAGTAGAGGAGCTTAGTAGGACGTCTTGCT	6623
QY	605	GCTGATGTCTCTACATATGACAGGCGGTCTGATACCGCCATATGGCGTGGAGAGAAACA	664
Db	6630	GCTGATGTCTCTACATATGACAGGCGGTCTGATACCGCCATATGGCGTGGAGAGAAACA	6685
QY	665	AGCTGCCATCAAGCGGTTGAGCAACTTTGTGCTGCTCACCAACAATGCTTACGCTA	724
Db	6690	AGCTGCCATCAATATGACTAGCAACTTTGTGCTGCTCACCAACAATGCTTATGCTA	6749
QY	725	CCACATCCCGAGCGCAAGCCAGCGGCGAGAAAGTCACTTTTGACAGACTGCMAATCC	784
Db	6750	CAACATCTCGAGCGCAAGCCAGCGGCGAGAAAGTCACTTTTGACAGACTGCAGGTCC	6805
QY	785	TGAGATGATCACTACAGGACGTGCTCAAGAGATATGAAGGCGAAGCGTCCACGTTAAG	844
Db	6810	TGAGATGATCACTACCGGACGTGCTCAAGAGATATGAAGGCGAAGCGTCCACGTTAAG	6865
QY	845	CTAAGCTTCTATCGTAGAGGACGTCGCAAGCTGACGCCCCCACTTTCGAGCAAACTA	904
Db	6870	CTAATCTTCTATCGTAGAGGACGCTGTATGACTGACGCCCCCACTTTCGAGCAAGTCA	6928
QY	905	AATTTGGCTATGGGCAAAAGACGTCGGAACCTATCCAGCAAGGCCATTAAACAATCC	964
Db	6930	AATTTGGCTATGGGCAAAAGACGTCGGAACCTATCCAGCAAGGCCATTAAACAATCC	6985
QY	965	GCTCCGTGTGGAGGACTTGTGGAGACACTGAACAACCAATTGACACCAATCATGG	1024
Db	6990	GCTCCGTGTGGAGGACTTGTCTGAGACACTGAGACCAATTAATGACACCAATCATGG	7049
QY	1025	CAAAAAATGAGGTTTTCTGCGTCCAACAGAGAGAGGCGCGGAAGCGACCTGGCCCTTA	1084
Db	7050	CAAAAAATGAGGTTTTCTGCGTCCAACAGAGAGGCGGCGCGAAGCGACCTGGCCCTTA	7105
QY	1085	TGCGTTCCTCAGACTTGTGGGGGTCCGTGTGTGTCAGAAATATGGCCCTCATAGCTGTGCT	1144

Db 7110 TCGTATTCACAGATTGGGGGTTCTGTGTGCGAGAAATGGCCCTTTAGAGTGTCT 7169
Qy 1145 CCAACCTCCCTCAGCGCTGATGGGCTCTCTGTAACGATTCAGATTCCTCTGAGACG 1204
Db 7170 CCAACCTCCCTCAGCGCGTATGGGCTCTCATACGATTCAGATTCCTCTGAGACG 7229
Qy 1205 GGGTCAGTTCCTGTAAGCGCTGAAATCAAGAAAGAGCCCTATGAGGCTTGCAATG 1264
Db 7230 GGGTCAGTTCCTGTAAGCGCTGAAATCAAGAAAGAGCCCTATGAGGCTTGCAATG 7289
Qy 1265 ACAACCGCTGTTTGAATCAAGTCACTGAGATGACATCCGTTAGAGATCAATTT 1324
Db 7290 ACAACCGCTGTTTGAATCAAGTCACTGAGATGACATCCGTTAGAGATCAATCT 7349
Qy 1325 ATGATGTTGTGATCTTGCGCCCGGAAAGCAGACAGCCATAGGTCCTCAGAGAGCG 1384
Db 7350 ACCAATGTTGTGATCTTGCGCCCGGAAAGCAGACAGCCATAGGTCCTCAGAGAGCG 7409
Qy 1385 TTTATATCGGGGCTCCCTGATCAATCAAAAGGAGAACTGCGGCTATCGCCGCTGCC 1444
Db 7410 TTTATATCGGGGCTCCCTGATCAATCAAAAGGAGAACTGCGGCTATCGCCGCTGCC 7469
Qy 1445 GCGGAGCGGCTGCTGACGACTAGCTGCGGTATACCTCAGATGTTACTTGAAGGCTT 1504
Db 7470 GCGGAGCGGCTGCTGACGACTAGCTGCGGTATACCTCAGATGTTACTTGAAGGCTT 7529
Qy 1505 CTGACAGCTGTGAGCTGCAAGAGCTCCAGAGCTGCAAGATGCTGCTGTCGAGAGCG 1564
Db 7530 CTGACAGCTGTGAGCTGCAAGAGCTCCAGAGCTGCAAGATGCTGCTGTCGAGAGCG 7589
Qy 1565 TTGTCCTTATCTGTGAGAGCGCGGAAACCAAGAGAGCGGCGGAGCTTACGATCTTCA 1624
Db 7590 TTGTCCTTATCTGTGAGAGCGCGGAAACCAAGAGAGCGGCGGAGCTTACGATCTTCA 7649
Qy 1625 CCGAGGCTATGACTAGTACTGTGCCCCCGGAGAGCGCGGAGAGCTGAGATAGACG 1684
Db 7650 CCGAGGCTATGACTAGTACTGTGCCCCCGGAGAGCGCGGAGAGCTGAGATAGACG 7709
Qy 1685 TGGAGTTGATACATCATGCTCTCCATGTTGTCGTCGCGAGAGATGCTTGGCAAAA 1744
Db 7710 TGGAGTTGATACATCATGCTCTCCATGTTGTCGTCGCGAGAGATGCTTGGCAAAA 7769
Qy 1745 GGGTATACCTACCTCAGCCGAGACC 1769
Db 7770 GGGTATACCTACCTCAGCCGAGACC 7794

RESULT 6
US-09-539-601-10
Sequence 10, Application US/09539601C
Patent No. 6630343
GENERAL INFORMATION:
APPLICANT: Bartschlagel, Ralf FW
TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
FILE REFERENCE: all sequences
CURRENT APPLICATION NUMBER: US/09/539,601C
EARLIER FILING DATE: 2001-08-30
EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
NUMBER OF SEQ ID NOS: 51
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 10
LENGTH: 7989
TYPE: DNA
ORGANISM: Hepatitis C virus
FEATURE:
NAME/KEY: 5'UTR
LOCATION: (1)..(341)
OTHER INFORMATION: construct I377/NS3-3'/wt
NAME/KEY: CDS
LOCATION: (342)..(1181)

OTHER INFORMATION: hepatitis C virus core-neomycin phosphotransferase
OTHER INFORMATION: fusion protein
FEATURE:
NAME/KEY: RBS
LOCATION: (1190)..(1800)
OTHER INFORMATION: internal ribosome entry site from
OTHER INFORMATION: encephalomyocarditis virus
FEATURE:
NAME/KEY: CDS
LOCATION: (1801)..(7758)
OTHER INFORMATION: hepatitis C virus NS3 - 5B
FEATURE:
NAME/KEY: 3'UTR
LOCATION: (7759)..(7989)
PUBLICATION INFORMATION:
AUTHORS: Lohmann, Volker
AUTHORS: Krieger, Frank
AUTHORS: Koch, Jan-Oliver
AUTHORS: Herian, Ulrike
AUTHORS: Theilmann, Lorenz
AUTHORS: Bartschlagel, Ralf
TITLE: Replication of subgenomic hepatitis C virus RNAs in a
TITLE: hepatoma cell line
JOURNAL: Science
VOLUME: 285
PAGES: 110-113
DATE: 1999-07-02
US-09-539-601-10

Query Match 87.7%; Score 1552.2; DB 4; Length 7989;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1632; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

Qy 5 ATGACTTCCGACGCGTGAAGCTTATGAGAGCAAGCTCTGTGCGGAGATGAGATGGCG 64
Db 5381 ATGACTTCCGACGCGTGAAGCTTATGAGAGCAAGCTCTGTGCGGAGATGAGATGGCG 5440
Qy 65 GGGACATTAACCGCGTGTGAGTCAAGAAACAGATTAATCTTGTGACCCCG 124
Db 5441 GGGACATTAACCGCGTGTGAGTCAAGAAACAGATTAATCTTGTGACCCCG 5500
Qy 125 TCCGAGCGAGAGAGATGAGCGGAGATGTCGTCGCGGAGATCTGCGGAAATCCA 184
Db 5501 TCCGAGCGAGAGAGATGAGCGGAGATGTCGTCGCGGAGATCTGCGGAAATCCA 5560
Qy 185 AGAATTTCCACACGAGTACCGGATGAGGAGAGCGCGGATTAACAACCTCCGCTGCG 244
Db 5561 AGAATTTCCACACGAGTACCGGATGAGGAGAGCGCGGATTAACAACCTCCGCTGCG 5620
Qy 245 AGTCTGGAAGAGCGCGGAGTACGTCCTCCAGTGTATCATGGGTCCCACTGCCACTTA 304
Db 5621 AGTCTGGAAGAGCGCGGAGTACGTCCTCCAGTGTATCATGGGTCCCACTGCCACTTA 5680
Qy 305 CTAAAGACCCCTCTATACCACTCCACGAGAAAGAGCAGTGTTCGACAGATCCA 364
Db 5681 CTAAAGACCCCTCTATACCACTCCACGAGAAAGAGCAGTGTTCGACAGATCCA 5740
Qy 365 CCGTGTCTTCTGCGCTGAGAGCTTGCACAAAGGCTTTGTAGTACTCCGACCGTGG 424
Db 5741 CCGTGTCTTCTGCGCTGAGAGCTTGCACAAAGGCTTTGTAGTACTCCGACCGTGG 5800
Qy 425 CCGTGAACGCGGAGAGCAAGCGCCCTCTGACCAATCTTCCAGAGAGCGGAGAGAG 484
Db 5801 CCGTGAACGCGGAGAGCAAGCGCCCTCTGACCAATCTTCCAGAGAGCGGAGAGAG 5860
Qy 485 GATCTGAGTGTGATGATGATTCCTCCATGCGCCCTTGAAGGAGAGCGGAGAGAGCCCG 544
Db 5861 GATCTGAGTGTGATGATGATTCCTCCATGCGCCCTTGAAGGAGAGCGGAGAGAGCCCG 5920
Qy 545 ATCTGAGCAGCGGCTTGTGATCAACGATGAGAGAGCGGATGAGAGCGTGTGCT 604
Db 5921 ATCTGAGCAGCGGCTTGTGATCAACGATGAGAGAGCGGATGAGAGCGTGTGCT 5980

QY 605 GCTCGATGTCCTACATGAGACGAGCGCTCTGATCAAGCCATGCGCTGCGAGGAAGCA 664
DB 5981 GCTCGATGTCCTACATGAGACGAGCGCTCTGATCAAGCCATGCGCTGCGAGGAAGCA 6040
QY 665 AGCTGCCATCAACGCGTTGAGCACTTTTGTGCTGATCAACAACATGCTTACGCTA 724
DB 6041 AGCTGCCATCAACGCGTTGAGCACTTTTGTGCTGATCAACAACATGCTTACGCTA 6100
QY 725 CCACATCCCGGAGCGGAGCGGAGAGAGAGGTCACCTTTGACAGACTGCAATCC 784
DB 6101 CAACATCTTCGAGCGGAGCGGAGAGAGAGGTCACCTTTGACAGACTGCGAGTCC 6160
QY 785 TGAAGCATCACTACAGAGAGCGTCTCAAGAGATGAAGGCGAGGCGTCCACAGTTAAG 844
DB 6161 TGAAGCATCACTACAGAGAGCGTCTCAAGAGATGAAGGCGAGGCGTCCACAGTTAAG 6220
QY 845 CTAACTTCTATCATGAGAGAGGCTGCAAGCTGACGCGCCCACTTGGCCAAATCTA 904
DB 6221 CTAACTTCTATCCGAGAGAGGCTGCAAGCTGACGCGCCCACTTGGCCAGATCTA 6280
QY 905 AATTGGCTATGAGGAGAGAGCGTCCGAGACCTTACGAGAGGCGATTACCAATCC 964
DB 6281 AATTGGCTATGAGGAGAGAGCGTCCGAGACCTTACGAGAGGCGATTACCAATCC 6340
QY 965 GCTCCGCTGAGAGAGAGCTTGTGAGAGACATGAACCAATTGACACCAACATCATG 1024
DB 6341 GCTCCGCTGAGAGAGAGCTTGTGAGAGACATGAACCAATTGACACCAACATCATG 6400
QY 1025 CAAAAAATGAGGTTTCTGCGTCCAAACGAGAGAGAGGCGGCAAGCGTCCGCTTA 1084
DB 6401 CAAAAAATGAGGTTTCTGCGTCCAAACGAGAGAGAGGCGGCAAGCGTCCGCTTA 6460
QY 1085 TCGTGTCCCAAGCTTGGGAGTCCGCTGTGAGAGAGAGGCGCTTATGAGCGTGT 1144
DB 6461 TCGTGTCCCAAGCTTGGGAGTCCGCTGTGAGAGAGAGGCGCTTATGAGCGTGT 6520
QY 1145 CCACCTCCCTCAGGCGTGTGATGAGGCTCTCGTACGATTCAGATTCCTCGAGCAGC 1204
DB 6521 CCACCTCCCTCAGGCGTGTGATGAGGCTCTCGTACGATTCAGATTCCTCGAGCAGC 6580
QY 1205 GGGTGAAGTCTCTGAGAGAGCGCTGGAATCAAGAGAGCGCTTATGAGCGTGT 1264
DB 6581 GGGTGAAGTCTCTGAGAGAGCGCTGGAATCAAGAGAGCGCTTATGAGCGTGT 6640
QY 1265 ACACCGCGCTTTTGAAGTCAAGAGTCAAGAGATGAGATCCGCTGAGAGAGTCAATT 1324
DB 6641 ACACCGCGCTTTTGAAGTCAAGAGTCAAGAGATGAGATCCGCTGAGAGAGTCAATT 6700
QY 1325 ATCAATGTTGATGAGTGGCGCCCGAGAGCGAGCAAGCGCATTAAGTCTGACAGAGCGC 1384
DB 6701 ACCAATGTTGATGAGTGGCGCCCGAGAGCGAGCAAGCGCATTAAGTCTGACAGAGCGC 6760
QY 1385 TTTTATATGAGGAGTCCCTGATTAATTTAAAGGAGGAGAGTGGCGCTTATGCGCGTGC 1444
DB 6761 TTTTATATGAGGAGTCCCTGATTAATTTAAAGGAGGAGAGTGGCGCTTATGCGCGTGC 6820
QY 1445 GCGGAGCGGCGTGTGAGAGCTAGTGGAGTAAATACCTTCATGATTTAGAGGCGT 1504
DB 6821 GCGGAGCGGCGTGTGAGAGCTAGTGGAGTAAATACCTTCATGATTTAGAGGCGT 6880
QY 1505 CTGCAAGCTGTGAGCTGCAAGAGCTTCAGAGTGAAGTGTCTGTGCGAGAGCGGC 1564
DB 6881 CTGCAAGCTGTGAGCTGCAAGAGCTTCAGAGTGAAGTGTCTGTGCGAGAGCGGC 6940
QY 1565 TTGTGCTTATCTGTGAGAGCGGAGAGCTCAAGAGAGAGCGGCGAGCTTACAGATCTTCA 1624
DB 6941 TTGTGCTTATCTGTGAGAGCGGAGAGCTCAAGAGAGAGCGGCGAGCTTACAGATCTTCA 7000
QY 1625 CGGAGGCTATGAGTACTCTGAGCGCGCGGAGAGCGGCGGAGAGAGAGTGAAGC 1684
DB 7001 CGGAGGCTATGAGTACTCTGAGCGCGCGGAGAGCGGCGGAGAGAGTGAAGC 7060
QY 1685 TGAAGTGAATACATCATGCTCTCCCAATGTGTGGTGGCGAGAGATGATCTGGCAAAA 1744

DB 7061 TGAAGTGAATACATCATGCTCTCCAAATGTGTGAGTGGCGAGAGATGATCTGGCAAAA 7120
QY 1745 GGGTACTACTACCAAGCGGTGAGCC 1769
DB 7121 GGGTACTACTACCAAGCGGTGAGCC 7145
RESULT 7
US-09-539-601-7
; Sequence 7, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FM
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; EARLIER FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 8001
; TYPE: DNA
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(341)
; OTHER INFORMATION: construct 1389/NS3-3'/wt
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (342)..(1193)
; OTHER INFORMATION: hepatitis C virus core-neomycin
; OTHER INFORMATION: phosphotransferase fusion protein
; FEATURE:
; NAME/KEY: RBS
; LOCATION: (1202)..(1812)
; OTHER INFORMATION: internal ribosome entry site from
; OTHER INFORMATION: encephalomyocarditis virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1813)..(7770)
; OTHER INFORMATION: hepatitis C virus nonstructural proteins NS3-5B
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (7771)..(8001)
; PUBLICATION INFORMATION:
; AUTHORS: Lohmann, Volker
; AUTHORS: Krner, Frank
; AUTHORS: Koch, Jan-Oliver
; AUTHORS: Herian, Ulrike
; AUTHORS: Theilmann, Lorenz
; AUTHORS: Bartenschlager, Ralf
; TITLE: Replication of subgenomic hepatitis C virus RNAs in a
; TITLE: hepatoma cell line
; JOURNAL: Science
; VOLUME: 285
; PAGES: 110-113
; DATE: 1999-07-02
US-09-539-601-7
Query Match 87.7%; Score 1552.2; DB 4; Length 8001;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1632; Conservative 0; Mismatches 133; Indels 0; Gaps 0;
QY 5 ATGACTTCCAGAGCGTGAAGCTTCAATGAGAGCAAGTCTCTGAGCGGAGATGAGCG 64
DB 5393 ATGACTTCCAGAGCGTGAAGCTTCAATGAGAGCAAGTCTCTGAGCGGAGATGAGCG 5452
QY 65 GGAACATTAACCGGCTGAGAGTCAAGAGCAAGTATCTTGAAGCTTTTGAAGCGCC 124
DB 5453 GGAACATTAACCGGCTGAGAGTCAAGAGCAAGTATCTTGAAGCTTTTGAAGCGCC 5512

125 TCCGAGCGGAGGAGATGAGCGGGAAGTCCCTCCCGGAGATCTCTGGGAATCCA 184
185 AGAATTTCCCAACAGGATGATCCCGCATGGGACGCGCGATTAACAACCTCCGCTGTGG 244
5573 GGAATTTCCCTCGAGGATGATCCCATATGGGACGCGCGATTAACAACCTCCGCTGTGG 5632
245 AGTCTGGAAGGCGCGGAGTACGCTCCCAATGGGATGATGAGGATGAGGATGAGGAT 304
5633 AGTCTGGAAGGAGCGGAGTACGCTCCCAATGGGATGATGAGGATGAGGATGAGGAT 5692
305 CTAAAGACCCCTCTCTATACCACTTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 364
5693 CCAAGGCGCGCTCGAATACCACTTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5752
365 CCGTCTCTTCTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 424
5753 CCGTCTCTTCTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5812
425 CCGTCTCTTCTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 484
5813 CCGTCTCTTCTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5872
485 GATCTGACGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 544
5873 GATCTGACGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5932
545 ATCTGACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 604
5933 ATCTGACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5992
605 GCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 664
5993 GCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6052
665 AGTCTGACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 724
6053 AGTCTGACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6112
725 CCAATCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 784
6113 CCAATCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6172
785 TGAGCATCTACTACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 844
6173 TGAGCATCTACTACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6232
845 CTAACTCTCTATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 904
6233 CTAACTCTCTATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6292
905 AATTGGCTATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 964
6293 AATTGGCTATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6352
965 GCTCCGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1024
6353 GCTCCGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6412
1025 CAAAAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1084
6413 CAAAAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6472
1085 TCGTATTCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1144
6473 TCGTATTCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6532
1145 CCAACCTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1204
6533 CCAACCTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6592

1205 GGGTGAAGTCTCTGTAAGAGCGCTGTAATCAAGAGAGCCCTATGAGGCTTTCATATG 1264
6593 GGGTGAAGTCTCTGTAAGAGCGCTGTAATCAAGAGAGCCCTATGAGGCTTTCATATG 6652
1265 ACAACCGGCTGTTTGAAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAG 1324
6653 ACAACCGGCTGTTTGAAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAG 6712
1325 ATCAATGTTGTAAGTGGGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1384
6713 ACCAATGTTGTAAGTGGGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6772
1385 TTTATATCGGAGGAGTCCCTGACTAATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1444
6773 TTTATATCGGAGGAGTCCCTGACTAATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6832
1445 GCGGAGCGGAGTGTGTAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCA 1504
6833 GCGGAGCGGAGTGTGTAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCA 6892
1505 CTGACGCTGTGTAAGTGTGTAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAG 1564
6893 CTGACGCTGTGTAAGTGTGTAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAG 6952
1565 TTGCTGTAATCTGTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1624
6953 TTGCTGTAATCTGTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 7012
1625 CCGAGGCTATGACTAGTACTGACGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1684
7013 CCGAGGCTATGACTAGTACTGACGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 7072
1685 TGAGTGAATACATCATGCTCTCTCAATGTGTGCTGCGGAGGAGGAGGAGGAGGAGGAGGAG 1744
7073 TGAGTGAATACATCATGCTCTCTCAATGTGTGCTGCGGAGGAGGAGGAGGAGGAGGAGGAG 7132
1745 GGGTATATACCTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1769
7133 GGGTATATACCTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 7157

RESULT 8
US-09-539-601-4
Sequence 4, Application US/09539601C
Patent No. 6630343
GENERAL INFORMATION:
APPLICANT: Bartenschlager, Ralf FW
TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
FILE REFERENCE: all sequences
CURRENT APPLICATION NUMBER: US/09/539,601C
EARLIER FILING DATE: 2001-08-30
EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
NUMBER OF SEQ ID NOS: 51
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 8637
TYPE: DNA
ORGANISM: Hepatitis C virus
FEATURE:
NAME/KEY: 5'UTR
LOCATION: (1)..(341)
OTHER INFORMATION: construct I377/NS2-3'/'wt
FEATURE:
NAME/KEY: CDS
LOCATION: (342)..(1181)
OTHER INFORMATION: HCV core-neomycin phosphotransferase fusion
OTHER INFORMATION: protein
FEATURE:
NAME/KEY: RBS
LOCATION: (1190)..(1800)
OTHER INFORMATION: internal ribosome entry site from
OTHER INFORMATION: encephalomyocarditis virus

FEATURE:
NAME/KEY: CDS
LOCATION: (1801).. (8406)
OTHER INFORMATION: hepatitis C virus NS2 - 5B
FEATURE:
NAME/KEY: 3' UTR
LOCATION: (8407).. (8637)
PUBLICATION INFORMATION:
AUTHORS: Lohmann, Volker
AUTHORS: Krner, Frank
AUTHORS: Koch, Jan-Oliver
AUTHORS: Herian, Uirike
AUTHORS: Thelmann, Lorenz
AUTHORS: Batteischlager, Ralf
TITLE: Replication of subgenomic hepatitis c virus RNAs in a
JOURNAL: Science
VOLUME: 285
PAGES: 110-113
DATE: 1999-07-02
US-09-539-601-4

Query Match 87.7%; Score 1552.2; DB 4; Length 8637;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1632; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

QY 5 ATGACTTCCAGACGCTGACCTCATCGAGGCCAACCTCTGTGGCGGACATGAGATGGCG 64
DB 6029 ATGACTTCCAGACGCTGACCTCATCGAGGCCAACCTCTGTGGCGGACATGAGATGGCG 6088
QY 65 GGGACATTACCCGCGTGGAGTCAAGAACAGATGATTAATCTTGAATCTTTCGACCCGC 124
DB 6089 GGAATCATACCCGCGTGGAGTCAAGAACAGATGATTAATCTTGAATCTTTCGACCCGC 6148
QY 125 TCCGACCGGAGAGATGAGTGGAGGAGTGTCTCCGCGGAGATCTTCGCGGAATCCA 184
DB 6149 TCCGACCGGAGAGATGAGTGGAGGAGTGTCTCCGCGGAGATCTTCGCGGAATCCA 6208
QY 185 AGAATATCCACAGAGATGAGTGGAGGAGTGTCTCCGCGGAGATCTTCGCGGAATCCA 244
DB 6209 GGAATATCCACAGAGATGAGTGGAGGAGTGTCTCCGCGGAGATCTTCGCGGAATCCA 6268
QY 245 AGTCTGGAAGGCGCGGAGTGTCTTCAGTGTGATGATGATGATGATGATGATGATGAT 304
DB 6269 AGTCTGGAAGGCGCGGAGTGTCTTCAGTGTGATGATGATGATGATGATGATGATGAT 6328
QY 305 CTAAGACCCCTCTATACCACTCCACGAGAAAGAGACATGTTCTTGAACAGATCCA 364
DB 6329 CTAAGACCCCTCTATACCACTCCACGAGAAAGAGACATGTTCTTGAACAGATCCA 6388
QY 365 CCGTGTCTTCTGCGTGGAGGAGTGTCTTCAGTGTGATGATGATGATGATGATGATGAT 424
DB 6389 CCGTGTCTTCTGCGTGGAGGAGTGTCTTCAGTGTGATGATGATGATGATGATGATGAT 6448
QY 425 CCGTGTCTTCTGCGTGGAGGAGTGTCTTCAGTGTGATGATGATGATGATGATGATGAT 484
DB 6449 CCGTGTCTTCTGCGTGGAGGAGTGTCTTCAGTGTGATGATGATGATGATGATGATGAT 6508
QY 485 GATCTGACGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 544
DB 6509 GATCTGACGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6568
QY 545 ATCTGACGACGAGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 604
DB 6569 ATCTGACGACGAGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 6628
QY 605 GCTGATGATCTTACATGAGGAGGCTTGTGATGATGATGATGATGATGATGATGATGAT 664
DB 6629 GCTGATGATCTTACATGAGGAGGCTTGTGATGATGATGATGATGATGATGATGATGAT 6688
QY 665 AGCTGACCATCAAGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 724
DB 6689 AGCTGACCATCAAGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 748

QY 725 CCACATCCCGAGCGCAAGCAGCGCAGAGAGAGTCACTTTGACAGACTGCATATCC 784
DB 6749 CAACATCTCGAGCGCAAGCCTGCGGAGAGAGAGTCACTTTGACAGACTGCAGTCC 6808
QY 785 TGGACGATCATCAAGAGAGTGTCTCAAGAGATGAGAGGCGGACCTTCAAGTAAAG 844
DB 6809 TGGACGATCATCAAGAGAGTGTCTCAAGAGATGAGAGGCGGACCTTCAAGTAAAG 6868
QY 845 CTAAGCTTCTATGATGAGAGAGCTGAGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 904
DB 6869 CTAAGCTTCTATGATGAGAGAGCTGAGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 6928
QY 905 AATTTGCTATGAGGAG 964
DB 6929 AATTTGCTATGAGGAG 6988
QY 965 GCTCCGTGTGGAG 1024
DB 6989 GCTCCGTGTGGAG 7048
QY 1025 CAAAAATGAGGTTTCTGCGTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1084
DB 7049 CAAAAATGAGGTTTCTGCGTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7108
QY 1085 TCGTGTCCAGAGCTTGGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1144
DB 7109 TCGTGTCCAGAGCTTGGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7168
QY 1145 CCACCTTCTCAGAGCTGTGATGAGGCTTCTGATGAGATGATGATGATGATGATGATGAT 1204
DB 7169 CCACCTTCTCAGAGCTGTGATGAGGCTTCTGATGAGATGATGATGATGATGATGATGAT 7228
QY 1205 GGGTGAATTCCTGTGAG 1264
DB 7229 GGGTGAATTCCTGTGAG 7288
QY 1265 ACACCGCTGTGTGATCAACAGTCACTGAGATGATGATGATGATGATGATGATGATGAT 1324
DB 7289 ACACCGCTGTGTGATCAACAGTCACTGAGATGATGATGATGATGATGATGATGATGAT 7348
QY 1325 ATCAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1384
DB 7349 ATCAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7408
QY 1385 TTATATGAGGAGGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1444
DB 7409 TTATATGAGGAGGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 7468
QY 1445 GCGGAGGAGGCTGTGAG 1504
DB 7469 GCGGAGGAGGCTGTGAG 7528
QY 1505 CTGACGCTGTGAG 1564
DB 7529 CTGACGCTGTGAG 7588
QY 1565 TTGTCTGATGATGAG 1624
DB 7589 TTGTCTGATGATGAG 7648
QY 1625 CCGAGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1684
DB 7649 CCGAGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7708
QY 1685 TGGAGTTGATTAATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1744
DB 7709 TGGAGTTGATTAATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7768
QY 1745 GGGTATATATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1769
DB 7769 GGGTATATATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 7793

```
RESULT 9
US-10-029-907-6
/ Sequence 6, Application US/10029907
/ Patent No. 6706874
/ GENERAL INFORMATION:
/ APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
/ TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
/ FILE REFERENCE: HEPATITIS C VIRUS
/ CURRENT APPLICATION NUMBER: US/10/029,907
/ PRIOR FILING DATE: 2001-12-21
/ PRIOR APPLICATION NUMBER: 60/257,857
/ PRIOR FILING DATE: 2000-12-22
/ NUMBER OF SEQ ID NOS: 25
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 6
/ LENGTH: 8638
/ TYPE: DNA
/ ORGANISM: HCV
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1802) ... (8407)
US-10-029-907-6
```

```
Query Match      87.7%; Score 1552.2; DB 4; Length 8638;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1632; Conservative 0; Mismatches 133; Indels 0; Gaps 0;
```

```
QY 5 ATGACTTCCGAGCGCTGACCTCATCGAGGCCAACCTCTGTGGCGCATGAGATGGCG 64
DB 6030 ATGACTTCCGAGCGCTGACCTCATCGAGGCCAACCTCTGTGGCGCATGAGATGGCG 6089
QY 65 GGAACATTACCCCGCTGAGTCAAGAAACAAGTATGTTCTTGGACTCTTTGACCCGC 124
DB 6090 GGAACATTACCCCGCTGAGTCAAGAAACAAGTATGTTCTTGGACTCTTTGACCCGC 6149
QY 125 TCGGACCGGAGAGATGAGCGGGAAGTGTCCCTCCCGGAGATCTCCGGAAATCCA 184
DB 6150 TCGGACCGGAGAGATGAGCGGGAAGTGTCCCTCCCGGAGATCTCCGGAAATCCA 6209
QY 185 AGAATTTCCACAGGAGATCCCGCATGGGACGCCCGATTAACAACCTCCGCTGTGG 244
DB 6210 GGAATTTCCACAGGAGATCCCGCATGGGACGCCCGATTAACAACCTCCGCTGTGG 6269
QY 245 AGTCTGGAAGGCCCGGAGTACGTCCTTCAGTGTGTAATGAGGTGCCAATGCCACTTA 304
DB 6270 AGTCTGGAAGGCCCGGAGTACGTCCTTCAGTGTGTAATGAGGTGCCAATGCCACTTA 6329
QY 305 CTAAAGACCCCTCTTATACCACTTCCACGAGAAAGAGACATGTTTCTGAACGAATCCA 364
DB 6330 CTAAAGACCCCTCTTATACCACTTCCACGAGAAAGAGACATGTTTCTGAACGAATCCA 6389
QY 365 CCGTGTCTTCTGCTGCGAGACTTGCCCAAAAGCTTTGGTAGCTCCGACCGTGG 424
DB 6390 CCGTGTCTTCTGCTGCGAGACTTGCCCAAAAGCTTTGGTAGCTCCGACCGTGG 6449
QY 425 CCGTGAACAGCGGACGCGCAACCGCCCTTCCTGACCAATCTTCGACGAGCGGAGCAG 484
DB 6450 CCGTGAACAGCGGACGCGCAACCGCCCTTCCTGACCAATCTTCGACGAGCGGAGCAG 6509
QY 485 GATCTGACGTTGATGATATTTCTCATGCCCCCTTGAGGGGGAGCCGGGGAGACCCCG 544
DB 6510 GATCTGACGTTGATGATATTTCTCATGCCCCCTTGAGGGGGAGCCGGGGAGATCCCG 6569
QY 545 ATCTCAGCAGCGGCTTTGGTCTACCGTGAATGAGAGCGCGGTGAGAGCGTGTGCT 604
DB 6570 ATCTCAGCAGCGGCTTTGGTCTACCGTGAATGAGAGCGCGGTGAGAGCGTGTGCT 6629
QY 605 GCTGAGTGTCTACATGAGAGAGCGGCTTGTATCAGCGCATGCGCTGCGAGAGAAACA 664
DB 6630 GCTGAGTGTCTACATGAGAGAGCGGCTTGTATCAGCGCATGCGCTGCGAGAGAAACA 6689
```

```
QY 665 AGTGCCCATCAACGGCTTGAAGCACTTTGTGCTGCTCACCAACAATGATCTACGCTA 724
DB 6690 AGTGCCCATCAACGGCTTGAAGCACTTTGTGCTGCTCACCAACAATGATGCTA 6749
QY 725 CCAATCTCCGAGCGGACGAGCGGACAGAAAGGTCACTTTGACAGACTGCAAAATCC 784
DB 6750 CCAATCTCCGAGCGGACGAGCGGACAGAAAGGTCACTTTGACAGACTGCAAGTCC 6809
QY 785 TGAAGCATCACTAACAGAGCGTGTCTAAGAGATGAAGCGGAGGCTCCACAGTTAAG 844
DB 6810 TGAAGCATCACTAACAGAGCGTGTCTAAGAGATGAAGCGGAGGCTCCACAGTTAAG 6869
QY 845 CTAACTCTATCTAGTGAAGAGAGCGTGAAGTGAAGCGCCCACTTGGCGCAATCTTA 904
DB 6870 CTAACTCTATCTAGTGAAGAGAGCGTGAAGTGAAGCGCCCACTTGGCGCAATCTTA 6929
QY 905 AATTGGCTATGGGAGAAAGAGAGCGTCCGAACTATCCAGAAAGCCATTAAACCATCC 964
DB 6930 AATTGGCTATGGGAGAAAGAGAGCGTCCGAACTATCCAGAAAGCCATTAAACCATCC 6989
QY 965 GCTCCGTGTGAGAGAGCTTTGGAAGACACTGAAACACCAATTGACACCATCATATG 1024
DB 6990 GCTCCGTGTGAGAGAGCTTTGGAAGACACTGAAACACCAATTGACACCATCATATG 7049
QY 1025 CAAAAAATGAGGTTTCTGCGTCCAAACAGAGAGAGAGCGGCAAGCCAGCTGCGCTTA 1084
DB 7050 CAAAAAATGAGGTTTCTGCGTCCAAACAGAGAGAGAGCGGCAAGCCAGCTGCGCTTA 7109
QY 1085 TCGTGTTCAGAGCTTGGGAGTCCGTGTGCGAGAAATGAGCCCTTATGAGCGTGTCT 1144
DB 7110 TCGTGTTCAGAGCTTGGGAGTCCGTGTGCGAGAAATGAGCCCTTATGAGTGTCT 7169
QY 1145 CCACTCTCCCTCAGAGTGTGATGAGGCTCTCTGTAAGGATTTCCAGATTTCTCTGACAC 1204
DB 7170 CCACTCTCCCTCAGAGTGTGATGAGGCTCTCTGTAAGGATTTCCAGATTTCTCTGACAC 7229
QY 1205 GGGTGAAGTCCGTGGAACGCGCTGGAATCAAGAAAGAGCCCTATGAGCTTGTGATATG 1264
DB 7230 GGGTGAAGTCCGTGGAATGCGCTGGAAGCGAAGAAATGCGCTTATGAGCTTGTGATATG 7289
QY 1265 ACAACCGCTGTTTGACTCAACAGTCACTGAGATGACATCCGTGTGAGAGATCAATT 1324
DB 7290 ACAACCGCTGTTTGACTCAACAGTCACTGAGATGACATCCGTGTGAGAGATCAATT 7349
QY 1325 ATCAATGTTGTGACTTGGGCCCCGAAAGCCAGACAGGCTTAAGTGTCTCAAGAGCGGC 1384
DB 7350 ACCAATGTTGTGACTTGGGCCCCGAAAGCCAGACAGGCTTAAGTGTCTCAAGAGCGGC 7409
QY 1385 TTTATATGGGGGGTCCCTGACTAATTCAAAAGGGAGAACTGGCGCTATGCGCGGTGCC 1444
DB 7410 TTTATATGGGGGGTCCCTGACTAATTCTAAAGGGAGAACTGGCGCTATGCGCGGTGCC 7469
QY 1445 GCGGAGCGGCTGTGCTGACGACTAGTGGGTATATCCCTCAATTTACTTTGAAGCGCT 1504
DB 7470 GCGGAGCGGCTGTGCTGACGACTAGTGGGTATATCCCTCAATTTACTTTGAAGCGCG 7529
QY 1505 CTGACGCTGTGAGAGTGTCAAAAGCTTCAGAGACTGCAAGTGTCTGTTGCGAGAGCGGC 1564
DB 7530 CTGCGGCTGTGAGAGTGTGAAAGCTTCAGAGACTGCAAGTGTCTGTTGCGAGAGCGGC 7589
QY 1565 TTGTGTTATCTGTGAGAGCGGGGAAACCAAGAGAGAGCGGCGAGCTTACAGAGTCTTA 1624
DB 7590 TTGTGTTATCTGTGAAAGCGGGGAAACCAAGAGAGAGCGGCGAGCTTACAGAGCTTCA 7649
QY 1625 CCGAGGCTATGACTAGTACTTGTGCCCCCGGAGGACCGGCGCAACCAAGATTAAGACC 1684
DB 7650 CCGAGGCTATGACTAGTACTTGTGCCCCCGGAGGACCGGCGCAACCAAGATTAAGACT 7709
QY 1685 TGAAGTTAATCATATATCTCTCCAAATGTGTGCGTGGCGAGATGATCTTGCGAANA 1744
DB 7710 TGAAGTTAATCATATATCTCTCCAAATGTGTGCGTGGCGAGATGATCTTGCGAANA 7769
QY 1745 GGGTATATACCTCAACCGGTGACCC 1769
```

```
Db      7770 GGGTGACTATCTCACCCGCTGACCC 7794
|||||
RESULT 10
US-10-029-907-24
; Sequence 24, Application US/10029907
; Patent No. 6706874
; GENERAL INFORMATION:
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
; FILE REFERENCE: 13/083
; CURRENT APPLICATION NUMBER: US/10/029,907
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,857
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 8638
; TYPE: DNA
; ORGANISM: HCV
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1802) ... (8407)
US-10-029-907-24

Query Match      87.7%; Score 1552.2; DB 4; Length 8638;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1632; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

QY      5 ATGACTTCCCAAGCGTGAAGCTATCGAGGCCAACCTCTGTGGCGGATGAGATGAGCG 64
Db      6030 ATGACTCCCCGAGACGCTGACCTCATCGAGGCCAACCTCTGTGGCGGAGATGAGCG 6089
QY      65 GGAACATTACCCGCGTGAAGTCAAGAACAAAGTAGTAATCTGCACTTTTGCACCCG 124
Db      6090 GGAACATCACCGCGGTGAAGTCAAGAAATTAAGTAGTAATTTTGAAGCTTTTGCAGCGCG 6149
QY      125 TCCGAGCGGAGAGATGAGTGGGAAAGTCCGTCGCCGCGAGATCTCTGGCGAAATCCA 184
Db      6150 TCCAAACGAGAGAGATGAGAGGAAAGTATCCGTTCCGCGAGATCTCTGGCGAGTCCA 6209
QY      185 AGAATATCCCAACGAGATGCGCGCATGGGACAGCCCGGATTCAACCTCTCGCTGTGG 244
Db      6210 GGAATATCTCTGAGGAGATGCCATATGAGGACAGCCCGATTCAACCTCTGATGTAG 6269
QY      245 AGTCTGGAAGGCCCGGACTAGCTGCTCCAGTGGTACATGAGTGGCCCACTGCACTTA 304
Db      6270 AGTCTGGAAGGCCCGGACTAGCTGCTCCAGTGGTACATGAGTGGTCCATTGCGCGCTG 6329
QY      305 CTAAAGCCCCCTCTATACCACTCCACGGAGAAAGAGACATGTGTCTGAACAGATCCA 364
Db      6330 CCAAGCCCCCTCGAATACCACTCCACGGAGAAAGAGACGATGTGTCTGATCAATCTTA 6389
QY      365 CCGTGTCTTCTGCGCTGCGGAGCTTGCCACAAAGGCTTTTGTAGCTCCGAGCGGTGG 424
Db      6390 CCGTGTCTTCTGCGCTGCGGAGCTTGCCACAAAGCTTTGCGAGCTCTCGAATGTGG 6449
QY      425 CCGTGAACGCGGACGCGAACCGGCCCTCTGACCAATCTCTCGACGAGCGGAGCAG 484
Db      6450 CCGTGAACGCGGACGCGAACCGGCCCTCTGACCAATCTCTCGACGAGCGGAGCAGCAG 6509
QY      485 GATCTGAGCTTATGATGCTATTTCTCGATGCCCCCTTGAAGGGGAGCGCGGGAGCCCG 544
Db      6510 GATCTGAGCTTATGATGCTATTTCTCGATGCCCCCTTGAAGGGGAGCGCGGGAGATCCG 6569
QY      545 ATCTCAGCAGCGGATCTTGTCTACCGTAGTGAAGGACCGGTGAGACGTCGTCTGCT 604
Db      6570 ATCTCAGCAGCGGATCTTGTCTACCGTAGTGAAGGACCGGTGAGACGTCGTCTGCT 6629
QY      605 GCTCATGTCTTACATGAGACAGGCGCTGTGATCAAGCCATGCGCTGCGAGGAAAGCA 664
```

```
Db      6630 GCTCATGTCTTACATGAGACAGGCGCTGTGATCAAGCCATGCGCTCGAGGAAACCA 6689
QY      665 AGCTGCCATCAACGCGGTTGAGCAACTCTTTGCTCCGCTACCAACAATGGTCTACGCTA 724
Db      6690 AGCTGCCATCAATGACATGAGCAACTCTTTGCTCCGCTACCAACAATGGTCTACGCTA 6749
QY      725 CCACATCCCGAGCGCAAGCAGCGCGGAGAAAGGTCACTTTGACAGACTGCAAAATCC 784
Db      6750 CAACATCTCGAGCGCAAGCTCTGCGGAGAAAGGTCACTTTGACAGACTGCAAGTCC 6809
QY      785 TGGAGATCACTTACCAAGACGCTGCTCAAGAGATGAAGGCGAAGGCTCCACAGTTAAG 844
Db      6810 TGGAGACCACTTACCGGAGCGTGTCTCAAGAGATGAAGGCGAAGGCTCCACAGTTAAG 6869
QY      845 CTAAGCTTCTATTCATGATGAAGAGCGCTGCAAGCGCCCAATTCGCGGCAAAATCTA 904
Db      6870 CTAACCTTCTATTCGATGAAGAGCGCTGTAAAGCTGACCGCCCAATTCGCGGCAAGATCTA 6929
QY      905 AATTGGCTATGGGCAAGAGACGTCGGAACCTATCCAGCAAGGCCATTAAACCAATCC 964
Db      6930 AATTGGCTATGGGCAAGAGACGTCGGAACCTATCCAGCAAGGCCATTAAACCAATCC 6989
QY      965 GCTCCGTGTGAAGACCTTGTGAGAGACACTGAACCAATTTGACACCAATCATGG 1024
Db      6990 GCTCCGTGTGAAGACCTTGTGAGAGACACTGAACCAATTTGACACCAATCATATGG 7049
QY      1025 CAAAAAATGAGGTTTTCGCGCCCAACGAGAGAGAGGCGGCAAGCCAGCTGCGCCTTA 1084
Db      7050 CAAAAAATGAGGTTTTCGCGCCCAACGAGAGAGAGGCGGCAAGCCAGCTGCGCCTTA 7109
QY      1085 TCGTGTCCCAAGACTTGGGGGTCGCTGTGTGTCGAGAAATGAGCCCTCTATGACGTGCT 1144
Db      7110 TCGTATTCGCAAGATTGGGGGTTGCTGTGTGTCGAGAAATGAGCCCTTTACATGTGCT 7169
QY      1145 CCACCTCTCTGAGGCTGTGATGGGCTCTCTGTACGATTCGAGTATCTCTGACACGC 1204
Db      7170 CCACCTCTCTGAGGCTGTGATGGGCTCTCTGTACGATTCGAGTATCTCTGACACGC 7229
QY      1205 GGGTGAAGTTCGTGTGAAGCGCTGGAATTAAGAAAGAACCCCATGAGGCTTGCATATG 1264
Db      7230 GGGTGAAGTTCGTGTGAATCTGTGAAGGGAAGAAATGCGCTATGAGGCTTGCATATG 7289
QY      1265 AACCCGCTGTTTGAATCAACAGTCACTGAGATGACATCCGTGAGAGAGTCAATTT 1324
Db      7290 AACCCGCTGTTTGAATCAACAGTCACTGAGATGACATCCGTGAGAGAGTCAATCT 7349
QY      1325 ATCAATGTTGACTTGGCCCCCGAAGCAGACAGGCCATTAAGTCCGCTCAAGAGCGGC 1384
Db      7350 ACCAATGTTGACTTGGCCCCCGAAGCAGACAGGCCATTAAGTCCGCTCAAGAGCGGC 7409
QY      1385 TTTATATGGGGGTCCTCCGCTGAATTCAAAGGCGAAGATGAGGCTATGCGGGTGC 1444
Db      7410 TTTATATGGGGGTCCTCCGCTGAATTCAAAGGCGAAGATGAGGCTATGCGGGTGC 7469
QY      1445 GCGGAGCGGCGTCTGACGACTAGCTGCGGTAAATACCTTCATATGTTACTTGAAGGCT 1504
Db      7470 GCGGAGCGGCGTCTGACGACTAGCTGCGGTAAATACCTTCATATGTTACTTGAAGGCG 7529
QY      1505 CTGACGCTGTGAGCTGCAAAAGCTCCAGACTCAGATGCTGTGTGCGGAGACGCGC 1564
Db      7530 CTGCGGCTGTGAGCTGCAAAAGCTCCAGACTCAGATGCTGTGTGCGGAGACGAGC 7589
QY      1565 TTGTGCTATCTGAGAGCGCGGGAACCCAGAGAGAGCGGCGAGCTTACAGATCTTCA 1624
Db      7590 TTGTGCTATCTGAGAGCGCGGGAACCCAGAGAGAGCGGCGAGCTTACAGATCTTCA 7649
QY      1625 CGAGGCTATGACTAGTACTCTGCCCCCGGAGGACCCGCCCAACAGAAATACGACC 1684
Db      7650 CGAGGCTATGACTAGTACTCTGCCCCCGGAGGACCCGCCCAACAGAAATACGACT 7709
QY      1685 TGAAGTTGATTAATCATGCTCTCTCAATGTGTGCGGTGCGACGATCATCTGAGCAAA 1744
```


Oy	1145	CCACCCTCCCTCAGGCTGTGATGGGCTCTCTGTACGGATTCGAGATTTCTCTGACAGC	1204
Db	7181	CCACCCTCCCTCAGGCTGTGATGGGCTCTCTGTACGGATTTCTCTGACAGC	7240
Oy	1205	GGGTGCAATTCCTGTGTGAACGCTTGAATCAAGAAACACCCCTATGGGCTTTGCATATG	1264
Db	7241	GGGTGCAATTCCTGTGTGAATGCTGTGAAGGAAGAAATGCTATGGGCTTTGCATATG	7300
Oy	1265	ACACCCGCTGTTTGTGACTTCACAGTCACTGAGATGACATCCGTGTAGGAGTCAATT	1324
Db	7301	ACACCCGCTGTTTGTGACTTCACAGTCACTGAGATGACATCCGTGTAGGAGTCAATT	7360
Oy	1325	ATCATATGTTGTGACTTTGGCCCCCGGAGCAGACAGGCGCATTAAGTGTGCTCACAGAGCGCC	1384
Db	7361	ACCATATGTTGTGACTTTGGCCCCCGGAGCAGACAGGCGCATTAAGTGTGCTCACAGAGCGCC	7420
Oy	1385	TTTATATCGAGGGGTCCTCTGACTTAATTCAAAAGGCGAGACTGCGGCTATCGCCGTC	1444
Db	7421	TTTACATCGAGGGGTCCTCTGACTTAATTCAAAAGGCGAGACTGCGGCTATCGCCGTC	7480
Oy	1445	GCGGAGCGCGGTGTGTACGACTGACTGCGTAATACCTTCAATGTTACTTTGAAGCCT	1504
Db	7481	GCGGAGCGCGGTGTGTACGACTGACTGCGTAATACCTTCAATGTTACTTTGAAGCCT	7540
Oy	1505	CTGCAGCCTGTGAGCTGCAAGAGCTCCAGAGCTGCGAGATGCTGTGTGGGAGAGCGCC	1564
Db	7541	CTGCAGCCTGTGAGCTGCAAGAGCTCCAGAGCTGCGAGATGCTGTGTGGGAGAGCGCC	7600
Oy	1565	TTGTGCTTATCTGTGAGAGCGCGGAAACCCAGAGAGAGCGCGGAGAGCTTACGAGTCTTCA	1624
Db	7601	TTGTGCTTATCTGTGAGAAAGCGCGGGAACCCAGAGAGAGCGAGGCGAGCTTACGAGTCTTCA	7660
Oy	1625	CGGAGGCTATGACTGATGATCTTGTGCCCCCCCCGGGAGACCCGCCCCACAGAAATTCGACC	1684
Db	7661	CGGAGGCTATGACTGATGATCTTGTGCCCCCCCCGGGAGACCCGCCCCACAGAAATTCGACT	7720
Oy	1685	TGAGATTGATTAACATCATGCTCTCTCCATATGTGTGCGTGGCGCAGATGCATCTGGCAAAA	1744
Db	7721	TGAGATTGATTAACATCATGCTCTCTCCATATGTGTGCGTGGCGCAGATGCATCTGGCAAAA	7780
Oy	1745	GGGTATACTACCTCACCCGTCGACCC	1769
Db	7781	GGGTATACTATCTCACCCGTCGACCC	7805
RESULT 14			
US-09-539-601-1			
; Sequence 1, Application US/09539601C			
; Patent No. 6630343			
; GENERAL INFORMATION:			
; APPLICANT: Bartschlagel, Ralf FW			
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System			
; FILE REFERENCE: all sequences			
; CURRENT APPLICATION NUMBER: US/09/539,601C			
; EARLIER FILING DATE: 2001-08-30			
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY			
; NUMBER OF SEQ ID NOS: 51			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 1			
; LENGTH: 11076			
; TYPE: DNA			
; ORGANISM: Hepatitis C virus			
; FEATURE:			
; NAME/KEY: 5'UTR			
; LOCATION: (1)..(341)			
; OTHER INFORMATION: construct 1389/Core-3'/wt			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (342)..(1193)			
; OTHER INFORMATION: hepatitis C virus core - neomycin			
; OTHER INFORMATION: phosphotransferase fusion protein			

	FEATURE:	NAME/KEY:	RBS	LOCATION:	(1202) .. (1812)	OTHER INFORMATION:	internal ribosome entry site from	OTHER INFORMATION:	encephalomyocarditis virus	FEATURE:	NAME/KEY:	CDS	LOCATION:	(1813) .. (10845)	OTHER INFORMATION:	hepatitis C virus polyprotein from core to	OTHER INFORMATION:	nonstructural protein NS5b; parental sequence	OTHER INFORMATION:	without cell culture-adaptive mutations	FEATURE:	NAME/KEY:	3'UTR	LOCATION:	(10846) .. (11076)	US-09-539-601-1
Query Match	87.7%	Score	1552.2	DB	4	Length	11076																			
Best Local Similarity	92.5%	Prod. No.	0	Mismatches	133	Indels	0	Gaps	0																	
Matches	1632	Conservative	0																							
Db	5	ATGACTTCCGAGCGTTCATCTCATCGAGGCGCAACTCTCTGTGGCGGATGATGGAGG	64																							
Db	8468	ATGACTCCCGGAGCGCTGACCTCATCTCAGGCGCAACTCTCTGTGGCGGAGATGGAGG	852																							
Qy	65	GGAACATTACCCGCGTGGAGTTCAGAAACANAGTATGTAATCCTTGACTCTTTGACCCGC	124																							
Db	8528	GGAACATCACCCGCGTGGAGTTCAGAAATAAGTATGTAATTTTGGACTCTTTGACGCCGC	858																							
Qy	125	TCCGAGGCGGAGAGATAGGCGGGAGAGTCCGTCGCCGCGGAGATCCTGGGAGAAATCCA	184																							
Db	8588	TCCAGCGGAGAGAGATAGAGAGAGATATCCGTCGCCGCGGAGATCCTGGGAGAGTCCA	864																							
Qy	185	AGAAATTCACACGAGATGCCGCGATGGGACGCGCCGAGATTACACCTCCGCTGCTGG	244																							
Db	8648	GGAATTCCTCCGAGCGATGCCCATATGGGACGCGCGGATTACACCTCCATGTTAG	870																							
Qy	245	AGTCTTGGAAAGCCCCCGGACTTACGTCTCTCCAGTGTATCATGGGGTCCCATCTGCACTTA	304																							
Db	8708	AGTCTTGGAAAGCCCCCGGACTTACGTCTCTCCAGTGTATCATGGGGTCTCATTTGCCCTG	876																							
Qy	305	CTAAGACCCCTCTATACCACTTCCAGAGAAAGAGAGAGTGTCTGACAGAAATCCA	364																							
Db	8768	CCAAGGCCCTCCGATACCACTTCCAGAGAAAGAGAGCGTTTCTGTACAGATCTA	882																							
Qy	365	CCGTGTCTTCTGCGCTTGGCGGAGCTTGCACAAAGCTTTTGTATGCTTCCGACCGTGG	424																							
Db	8828	CCGTGTCTTCTGCGCTTGGCGGAGCTTGCACAAAGCTTGGCGACGCTCCGATATGTGG	888																							
Qy	425	CCGTGACAGGAGGACGCGCAACGCGCCCTCCGACCAATCTCCGACGACGCGCGAGAG	484																							
Db	8888	CCGTGACAGGAGGACGCGCAACGCGCTTCTGACACGCTTCCGACGACGCGCGAG	894																							
Qy	485	GATCTGACGTTGAGTGTGTAATCTTCATGCCGCCCTTGAAGGGGAGCGCGGGACCCCG	544																							
Db	8948	GATCTGACGTTGAGTGTGTAATCTTCATGCCGCCCTTGAAGGGGAGCGCGGGATCCCG	900																							
Qy	545	ATCTCAGCGAGCGGCTTTTGTGTCTACCGTGAATGAGAGGCGCGTGTAGGACGTCTGTCT	604																							
Db	9008	ATCTCAGCGAGCGGCTTTTGTGTCTACCGTGAATGAGAGGCTGTGTGTCTGTCT	906																							
Qy	605	GCTGATGTCTTACACATGACAGGCGCTGTATCAGCGCCATGTGCGTTCGAGAGAAACA	664																							
Db	9068	GCTGATGTCTTACACATGACAGGCGCTGTATCAGCGCCATGTGCGTTCGAGAGAAACA	912																							
Qy	665	AGTGGCCCATCAACGCGTTGAGCAACTTTGTGCGTCAACACAACTGTGTCTAC																								

Db	9248	TTGACGACCACTACCGGAGCGTGTCAAGAGATGAGGCGAGGCGTCCACAGTTAAG	9307
Qy	845	CTAAGCTTCTATCAGTAGAGGAAGCTTGCAAGCTGACGCGCCCAATTCCGCCAAATCTA	904
Db	9308	CTAAACTTCTATCCGTGGAGGAGGCTGTAGCTGACGCCCCACATTCGGCCAGATCTA	9367
Qy	905	AAATTGGCTATGGGGGCAAAAGACGTCCTGGGAACCTATCCAGAAAGGCCATTAAACAATCC	964
Db	9368	AAATTGGCTATGGGGGCAAAAGACGTCCTGGGAACCTATCCAGAAAGGCCATTAAACAATCC	9427
Qy	965	GCTCCGTGTGGAGGAGACTTGTGTGAAGACCTGAACCACTAATTGACCAACCATCATG	1024
Db	9428	GCTCCGTGTGGAGGAGACTTGTGTGAAGACCTGAACCACTAATTGACCAACCATCATG	9487
Qy	1025	CAAAAAATGAGGTTTTTCTGCGTCCACCAAGAGAGAGAGGCGCCGACGCTGCGCTTA	1084
Db	9488	CAAAAAATGAGGTTTTTCTGCGTCCACCAAGAGAGAGAGGCGCCGACGCTGCGCTTA	9547
Qy	1085	TCGTGTTCCCAAGACTTGGGGGGTCGTGTGTGCGAAGAAAATGAGCCCTCATGACGTGCT	1144
Db	9548	TCGTATTCACAGATTTTGGGGGTTGCTGTGTGCGAAGAAAATGAGCCCTTATGATGTGCT	9607
Qy	1145	CCACCTTCCTCAAGGCTGTGATGGGCTCTCTGTACGATTCAGATTTCTCTGAGACG	1204
Db	9608	CCACCTTCCTCAAGGCGGTGATGGGCTCTTCAATCGGATTCCAATATCTCTTGAGACG	9667
Qy	1205	GGGTGCAATTCCTGGTGAAGGCTGGAAATCAAGAAAGAGCCCATAGGGGCTTTCATATG	1264
Db	9668	GGGTGCAATTCCTGGTGAATGCTGGAAAGCGAAAGAAATGCCCTTATGGGCTTTCGCAATG	9727
Qy	1265	ACACCCGCTGTTTGTGACTCAACAGTCACTGGAATGACATTCGTTGAGAGAGTCAATTT	1324
Db	9728	ACACCCGCTGTTTGTGACTCAACAGGTCATGGAATGACATTCGTTGAGAGTCAATTC	9787
Qy	1325	ATCATATGTTTGTGACTTGGGCCCCCGCAAGCAACAGGCCATTAAGTTCGCTCAAGAGGCG	1384
Db	9788	ACCAATGTTGTGACTTGGGCCCCCGCAAGCAACAGGCCCATTAAGTTCGCTCAAGAGGCG	9847
Qy	1385	TTTATATGCGGGGTCCTCTGACTAATTCAAAAGGGCAGAACTGGCGGCTATGGCGGTGC	1444
Db	9848	TTTATATGCGGGGTCCTCTGACTAATTCAAAAGGGCAGAACTGGCGGCTATGGCGGTGC	9907
Qy	1445	GCGGAGCGCGCGTGTGACGACTAGCTGCGGTAATACCTTCACATGTTACTTTGAAGGCT	1504
Db	9908	GCGGAGCGCGGTGTGACGACAGCTGCGGTAATACCTTCACATGTTACTTTGAAGGCG	9967
Qy	1505	CTGCAAGCTGTGAGAGCTCAAGAGCTCCAGAGCTGACGATGCTGTGGCGAGACGAGCC	1564
Db	9968	CTGCGGCGCTGTGAGAGCTCGAAGCTCCAGAGCTGACGATGCTGTGTGCGAGACGAGCC	1002
Qy	1565	TTGTGCTTATCTGTAGAGCGCGGAAACCAAGAGAGAGCGCGGCAAGCTTACGATCTTCA	1624
Db	10028	TTGTGCTTATCTGTAGAAAGCGCGGGAACCAAGAGAGAGCGGCGAAGCTTACGAGCTTCA	1008
Qy	1625	CGAGGCGCTATGACTAGGATCTGCGCCCCCTGGGGAGCCGCGCCAAACCAAAATACGACT	1684
Db	10088	CGAGGCGCTATGACTAGGATCTGCGCCCCCTGGGGAGCCGCGCCAAACCAAAATACGACT	1014
Qy	1685	TGAGATTGATAACATCATGCTCTTCCATATGTGTGCGTGCAGACGATGCATCTGGCAAA	1744
Db	10148	TGAGATTGATAACATCATGCTCTTCCATATGTGTGCGTGCAGACGATGCATCTGGCAAA	1020
Qy	1745	GGGTATATACCTCAACCGGTACCC	1769
Db	10208	GGGTATATATCTCAACCGGTACCC	10232

Db 6750 CAACATCTGACGCGCAAGCTTGCAGAGAAAGGTCACTTTGACAGACTGCAGGTCC 6809
QY 785 TGGAGATACATACCAAGACGCTGCTCAAGAGATGAAGGCGAAGGCTCCACAGTTAAG 844
Db 6810 TGGAGACCACTACCGGACGCTGCTCAAGAGATGAAGGCGAAGGCTCCACAGTTAAG 6869
QY 845 CTAGCTTCTATCAGTAGAGAAAGCTGCAAGCTGACGCCCCACATTGGCCAAATCTA 904
Db 6870 CTAAACTTCTATCCGTGAGAGAAAGCTGTAAGCTGACGCCCCACATTGGCCAGATCTA 6929
QY 905 AATTGGCTATGGGCAAAAGACGTCCGGAACCTATCCAGCAAGGCCATTAAACAATCC 964
Db 6930 AATTGGCTATGGGCAAAAGACGTCCGGAACCTATCCAGCAAGGCCGTTAAACAATCC 6989
QY 965 GTCCTGTTGGAGAGACTTTGGAAAGACCTGAAACACCAATTGACACCAACATCATAG 1024
Db 6990 GTCCTGTTGGAGAGACTTTGGAGAGACCTGAAAGACCAATTTGACACCAACATCATAG 7049
QY 1025 CAAAAAATGAGSTTTTCTGCTCCAAACAGAGAGAGAGGCGCAAGCCAGCTCGCTTA 1084
Db 7050 CAAAAAATGAGSTTTTCTGCTCCAAACAGAGAGAGAGGCGCGCAAGCCAGCTCGCTTA 7109
QY 1085 TCGTGTCCAGACTTGGGGGTCCTGCTGTCGAGAAATGGCCCTTAATGACGTGCT 1144
Db 7110 TCGTATTCAGATTTGGGGGTCCTGCTGTCGAGAAATGGCCCTTAATGATGTGCT 7169
QY 1145 CCAGCTCCCTCAGCTGTGATGGGCTCCTCGTACGATTCAGATTTCTCTGACAGC 1204
Db 7170 CCAGCTCCCTCAGCTGTGATGGGCTCCTCATACGATTCCAATATCTCTGACAGC 7229
QY 1205 GGGTGAAGTCTCGGTGAACGCTGGAATCAAGAAAGACCCCTAATGGCTTTGATATG 1264
Db 7230 GGGTGAAGTCTCGGTGAATGCTTGAAGAAAGCAAGAAATGCCCTATGGCTTCGATATG 7289
QY 1265 ACACCCGCTGTTTGACTCAACAGTCACTGAGATGACATCCGTGTAGAGAGTCAATT 1324
Db 7290 ACACCCGCTGTTTGACTCAACAGTCACTGAGATGACATCCGTGTAGAGAGTCAATCT 7349
QY 1325 ATCAATGTTGATCTTGGGCCCCGGAAGCCAGACAGGCCATAAGTTCCTCAAGAGCGGC 1384
Db 7350 ACCAATGTTGATCTTGGGCCCCGGAAGCCAGACAGGCCATAAGTTCCTCAAGAGCGGC 7409
QY 1385 TTTATATCGGGGGTCCCTGACTAATTCAAAAGGCGAAGCTGCGCTATCGCCGATGCC 1444
Db 7410 TTTATATCGGGGGTCCCTGACTAATTCAAAAGGCGAAGCTGCGCTATCGCCGATGCC 7469
QY 1445 GCGCAGCGGCTGTGACGACTAGTGGGTAATACCTCAATGTTACTTGAAGGCT 1504
Db 7470 GCGCAGCGGCTGTGACGACTAGTGGGTAATACCTCAATGTTACTTGAAGGCT 7529
QY 1505 CTGCAAGCTGTGAGCTGCAAGCTTCAGGACTGCAAGTGTGCTGTGCGGAGAGCGGC 1564
Db 7530 CTGCGGCTGTGAGCTGCAAGCTTCAGGACTGCAAGTGTGCTGTGCGGAGAGCGGC 7589
QY 1565 TTGTGCTTATCTGTGAGACGCGGGAACCTCAGAGAGAGCGGCGACTACGATGTTCA 1624
Db 7590 TTGTGCTTATCTGTGAGAGCGGGAACCTCAGAGAGAGCGGCGACTACGAGCTTCA 7649
QY 1625 CCGAGGCTATGACTAGTACTGTGCCCTCGGGGACCGGCCAACCAGAAATGACCT 1684
Db 7650 CCGAGGCTATGACTAGTACTGTGCCCTCGGGGACCGGCCAACCAGAAATGACCT 7709
QY 1685 TGGAGTTGATACATCATGCTCTCTCAATGTGTGCTGCGGACGATGCTTGGCAAAA 1744
Db 7710 TGGAGTTGATACATCATGCTCTCTCAATGTGTGCTGCGGACGATGCTTGGCAAAA 7769
QY 1745 GGGTATACTACTCAACCGGTGACC 1769
Db 7770 GGGTATACTACTCAACCGGTGACC 7794



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 144964

TO: Bao-Qun Li
Location: rem/3d24/3c18
Art Unit: 1648
Tuesday, February 22, 2005

Case Serial Number: 09/664363

From: Edward Hart
Location: Biotech-Chem Library
REM-1A55
Phone: 571-272-2512

edward.hart@uspto.gov

Search Notes

Examiner Li,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

This Page Blank (uspto)

3

STIC-Biotech/ChemLib

144964

my

From: Li, Bao-Qun
Sent: Friday, February 11, 2005 8:12 AM
To: STIC-Biotech/ChemLib

PLEASE DO THE SEQUENCE HOMOLOGY AND INTERFERENCE OF BASES 308-2116 OF SEQ ID NO; 21 IN APPLICATION SN. 09,664,363. THANKS.

Bao Qun Li M.D
TC 1600
Art Unit 1648
Tel. 517-272-0904
REM, 3C18
Rm. 3D24

RECEIVED
FEB 11 2005
STIC

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2- _____
Date Searcher Picked up: 2/16/05
Date Completed: 2/16/05
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA Sequence: # _____
AA Sequence: # _____
Structure: # _____
Bibliographic: _____
Litigation: _____
Patent Family: _____
Other: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: 256
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)



STIC SEARCH RESULTS

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact:*

Mary Hale, Information Branch Supervisor
Remsen Bldg. 01 D86
571-272-2507

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC-Biotech-Chem Library Remsen Bldg.



This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2005, 21:00:40 ; Search time 8042.09 Seconds
(without alignments)
10899.589 Million cell updates/sec

Title: US-09-664-363-21_COPY_308_2116

Perfect score: 1809
Sequence: 1 ATGAGCAGCAATCTTAACC.....GGCGTTGGGGCCTTGTTG 1809

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sbs:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1809	100.0	2116	6 A28155	A28155 PT-NANBH mr
2	1809	100.0	2116	6 A32202	A32202 NANBH PT P
3	1809	100.0	2116	6 AR144050	AR144050 Sequence
4	1605.8	88.8	9410	14 HPCX182	D50481 Hepatitis C
5	1593.6	88.1	9359	14 AF313916	AF313916 Hepatitis C
6	1593.6	88.1	9410	14 HPCX182	D50485 Hepatitis C
7	1581.8	87.4	9431	14 HCU45476	U45476 Hepatitis C
8	1581.8	87.4	9585	14 AB049095	AB049095 Hepatitis C
9	1580.2	87.4	9379	14 AF207761	AF207761 Hepatitis C
10	1580.2	87.4	9379	14 AF207766	AF207766 Hepatitis C
11	1575.4	87.1	9547	14 AB049091	AB049091 Hepatitis C
12	1572.2	86.9	9605	6 AX739971	AX739971 Sequence
13	1572.2	86.9	9605	14 HCU238799	HCU238799 Hepatitis C
14	1570.6	86.8	3286	14 AB008441	AB008441 Hepatitis C
15	1570.6	86.8	9033	14 HCU238800	HCU238800 Hepatitis C
16	1570.6	86.8	11076	6 AR406040	AR406040 Sequence
17	1570.6	86.8	11076	6 AR406046	AR406046 Sequence
18	1570.6	86.8	11076	6 AR406048	AR406048 Sequence
19	1570.6	86.8	11076	6 AR406050	AR406050 Sequence

20	1570.6	86.8	11076	6 AX036252	AX036252 Sequence
21	1570.6	86.8	11076	6 AX036258	AX036258 Sequence
22	1570.6	86.8	11076	6 AX036260	AX036260 Sequence
23	1570.6	86.8	11076	6 AX036262	AX036262 Sequence
24	1569.6	86.7	9580	14 AF054250	AF054250 Hepatitis C
25	1568.6	86.7	9410	14 HPCX181	D50480 Hepatitis C
26	1567.4	86.6	9377	14 AF207756	AF207756 Hepatitis C
27	1567.4	86.6	9460	14 HPCX191	D10750 Hepatitis C
28	1567.4	86.6	9533	14 HCU132997	HCU132997 Hepatitis C
29	1565.8	86.6	2540	6 E04260	E04260 CDNA encodi
30	1565.8	86.6	2540	6 E04805	E04805 CDNA to 5'-
31	1565.8	86.6	2540	6 E07391	E07391 CDNA encodi
32	1565.8	86.6	9448	14 HPCX183	D13558 Hepatitis C
33	1564.2	86.5	3296	14 AB008442	AB008442 Hepatitis C
34	1564.2	86.5	9456	14 HPCX184	D10934 Hepatitis C
35	1564.2	86.5	9598	14 AB049101	AB049101 Hepatitis C
36	1563.2	86.4	9379	14 AF165052	AF165052 Hepatitis C
37	1562.6	86.4	3296	14 AB008443	AB008443 Hepatitis C
38	1562.6	86.4	8781	14 AF054254	AF054254 Hepatitis C
39	1562.6	86.4	9375	14 AF207754	AF207754 Hepatitis C
40	1562.6	86.4	9379	14 AF165051	AF165051 Hepatitis C
41	1562.6	86.4	9596	14 AF054249	AF054249 Hepatitis C
42	1561	86.3	9573	14 AB049090	AB049090 Hepatitis C
43	1560	86.2	9410	14 HPCX181	D50483 Hepatitis C
44	1559.4	86.2	8780	14 AF054259	AF054259 Hepatitis C
45	1559.4	86.2	9369	14 AF165053	AF165053 Hepatitis C

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	JOURNAL	FEATURES	source	CDS	ORIGIN
A28155	A28155	PT-NANBH mRNA fragment from patent GB239245.	A28155	A28155.1	GI:1248638	synthetic construct	synthetic construct	other sequences; artificial sequences.						
								1 (bases 1 to 2116)						
								Post-translational non-A non-B hepatitis viral polypeptides						
								Patent: GB 229245-A 21-26-JUN-1991;						
								Location/Qualifiers						
								1..2116						
								/organism="synthetic construct"						
								/mol_type="unassigned DNA"						
								/db_xref="taxon:32630"						
								/note="contig formed by CDNA clones"						
								308..>2116						
								/codon_start=1						
								/transl_table=1						
								/product="PT-NANBH polypeptide N-terminus"						
								/protein_id="CAA01931.1"						
								/db_xref="GI:1248639"						
								/translation="MSTNPKPQKTKRNTNRRPQDKPEGGQIVGVYLLRRGRRL						
								GVRAATKTSERSQPRGRQPIPKARQPEGRMAQGVPEPLVGNMGMAAGWLLSPRG						
								SRPWSGPTDPRRSRNILGVIDLTLCGFADLGVYPLVGAAPGAAALAAHLYVED						
								GVYATNGNPGCSFSIFLLALSCITPAAVEVNAVSGIYVTDNCSSSLVYEAD						
								MTMHPGVCPCVRENGSSRCWALPTPLAAKASLPTATIRPHVDLYNCAAFPCAMV						
								VDDLCGSFVLSQLPTFSRRRQTVQDCNCSLIPHVGSHRAMDMANMSEPTALAV						
								SQLRLPQAVVDVVAAGHGVLAGVAVYSWGNMAKVLVWMLFAGVDEPYTTGTH						
								GRHAAGLTSLETPGPAOKIOLVNTNGSHNTRTALNCNDSLTGRLAALFYTHRNAS						
								GCSERMASCRPIDPQDQMGPIYVNESGLDQRPVCMHAYAPQCGIVPALVQCVGYC						
								FTSPVAVGTTDRFGAPYTRMGENTDVLANNTRPVGNNMGCTMANNSTGPTKTCGG						
								PCNNIGVGNNTLICPTDCFRKHPRATYTKCGSPWL"						

Query Match 100.0%; Score 1809; DB 6; Length 2116;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1809; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGAGCAGCAATCTCTAAACCTCAAGAAAAACCAACGTAACCAACCCGCCCCACAG 60
308 ATGAGACGAATCTCTAAACCTCAAGAAAAACCAACGTAACCAACCCGCCCCACAG 367
61 GAGCTCAAGTTCCTCCGGGCGGTGTCAGATCGTTGTGGAGTTTACCTGTTGCCGCGAG 120
368 GAGCTCAAGTTCCTCCGGGCGGTGTCAGATCGTTGTGGAGTTTACCTGTTGCCGCGAG 427
121 GAGCCAGGTTGGGTGTGGCGCGAGCTAGGAAGACTTCGAGCGGTCCGAACCTGTGTGA 180
428 GAGCCAGGTTGGGTGTGGCGCGAGCTAGGAAGACTTCGAGCGGTCCGAACCTGTGTGA 487
181 AGGCGCAACACTATCCCAAGGCTCGCCAGCCGAGGGGAGGGGCTGAGCGCCGGG 240
488 AGGCGCAACACTATCCCAAGGCTCGCCAGCCGAGGGGAGGGGCTGAGCGCCGGG 547
241 TACCTTGGCCCTCTATGAGCAACGAGGGCATGAGGGTGGGAGAGATGCTCTGTCAACC 300
548 TACCTTGGCCCTCTATGAGCAACGAGGGCATGAGGGTGGGAGAGATGCTCTGTCAACC 607
301 CGTGGCTCCGGGCTAGTGGGGGCCCCATGACCCCGGCGTAAAGTCCGCTAATTGGGT 360
608 CGTGGCTCCGGGCTAGTGGGGGCCCCATGACCCCGGCGTAAAGTCCGCTAATTGGGT 667
361 AAAGTCATGATACCTCTCATGCGGCTTGCCGCACTTCATGAGGATCAATTCGCTCGTC 420
668 AAAGTCATGATACCTCTCATGCGGCTTGCCGCACTTCATGAGGATCAATTCGCTCGTC 727
421 GGGCTCCCTTAGGGGGGCGGTGCGAGGGCCCTGGCGCATGCGTCCGGGTTCTGAGAGAC 480
728 GGGCTCCCTTAGGGGGGCGGTGCGAGGGCCCTGGCGCATGCGTCCGGGTTCTGAGAGAC 787
481 GGGCGAATATGCAACAGGGAATTTACCGGGTGTCTCTTCTATCTTCTCTTGGT 540
788 GGGCGAATATGCAACAGGGAATTTACCGGGTGTCTCTTCTATCTTCTCTTGGT 847
541 TTGCTGCTCTGTTTGAACCATTCAGCTTCGCTTAATGAGTGCAGCACTGTCCGGATC 600
848 TTGCTGCTCTGTTTGAACCATTCAGCTTCGCTTAATGAGTGCAGCACTGTCCGGATC 907
601 TACCATGTACAGAACGATGTTCTCACTCAAGCATGCTGTACGAGACAGCGGACATGATC 660
908 TACCATGTACAGAACGATGTTCTCACTCAAGCATGCTGTACGAGACAGCGGACATGATC 967
661 ATGCAACCCCGGGGTGTGGCGGTGTCGCGGAGGGGAATTCCTCCCGCTGCTGGGTA 720
968 ATGCAACCCCGGGGTGTGGCGGTGTCGCGGAGGGGAATTCCTCCCGCTGCTGGGTA 1027
721 GGGCTCACTCCACAGCTCGCGGCAAGAGCGCAGCATCCCACTGCGACAAATACGACGC 780
1028 GGGCTCACTCCACAGCTCGCGGCAAGAGCGCAGCATCCCACTGCGACAAATACGACGC 1087
781 CAGCTGATTTGCTGTTGGGGGCGTGCCTTCTGCTCGCTATGATCGTGGGGGATCTC 840
1088 CAGCTGATTTGCTGTTGGGGGCGTGCCTTCTGCTCGCTATGATCGTGGGGGATCTC 1147
841 TGGGATCTGTTTCTGCTCTCTGATGCTGTCACTTTCGCTTCGCGCAATACGAGC 900
1148 TGGGATCTGTTTCTGCTCTCTGATGCTGTCACTTTCGCTTCGCGCAATACGAGC 1207
901 GTACAGAGATGCAATTTGTTCAATCTATCCCGGCAAGTATCAGGTCAACGATGCTGG 960
1208 GTACAGAGATGCAATTTGTTCAATCTATCCCGGCAAGTATCAGGTCAACGATGCTGG 1267
961 GATATGATATGAATCTGTACCTTACAGACAGCCCTATGATGATTCGAGCTACTCCGATC 1020
1268 GATATGATATGAATCTGTACCTTACAGACAGCCCTATGATGATTCGAGCTACTCCGATC 1327
1021 CCAAGAGCTGTGTGTGACATGTGTGGGGGGGCGCACTGGGGAGTCTGTGGCGGCTTGGC 1080
1328 CCAAGAGCTGTGTGTGACATGTGTGGGGGGGCGCACTGGGGAGTCTGTGGCGGCTTGGC 1387

1081 TACTATTCATGATGTGGGAACTGGGCTTAAGGTCTTGTGTGTGATGCTACTCTTTGCCGCGC 1140
1388 TACTATTCATGATGTGGGAACTGGGCTTAAGGTCTTGTGTGTGATGCTACTCTTTGCCGCGC 1447
1141 GTTGAAGGGGAACTTTAACAACAAGAGGGGGAACAACGGCGCGGCCCAAGGGCTTACA 1200
1448 GTTGAAGGGGAACTTTAACAACAAGAGGGGGAACAACGGCGCGGCCCAAGGGCTTACA 1507
1201 TCCCTCTTACACACCTGGGCGGGCTCAGAAAATCCAGCTTGTAAACCAACGAGGAGCTGG 1260
1508 TCCCTCTTACACACCTGGGCGGGCTCAGAAAATCCAGCTTGTAAACCAACGAGGAGCTGG 1567
1261 CACATCAACAGAACTGCTTGAATGCAATGACTCCCTCAAACTGGGTTCTTGGCGCG 1320
1568 CACATCAACAGAACTGCTTGAATGCAATGACTCCCTCAAACTGGGTTCTTGGCGCG 1627
1321 CTGTTCTACACGACAGGTTCAATGCGTCCGGAATGCTCAGAGCGCATGCGAGCTGCCG 1380
1628 CTGTTCTACACGACAGGTTCAATGCGTCCGGAATGCTCAGAGCGCATGCGAGCTGCCG 1687
1381 CCATTTGACAGATTCGATCAGAGGGGAGGTCCTCACTTAATTAATGAGTCCACGAGCTTG 1440
1688 CCATTTGACAGATTCGATCAGAGGGGAGGTCCTCACTTAATTAATGAGTCCACGAGCTTG 1747
1441 GACCAAGAGCCCTATATGCTGGCACTACGACCTCAACCGGTGGTATGTCGCCGCTTG 1500
1748 GACCAAGAGCCCTATATGCTGGCACTACGACCTCAACCGGTGGTATGTCGCCGCTTG 1807
1501 CAGGTGTGTGGCCCAAGTGTACTGTTTCACTCAAGCCCTGTTGTGTGTGGGAGCAACGAT 1560
1808 CAGGTGTGTGGCCCAAGTGTACTGTTTCACTCAAGCCCTGTTGTGTGTGGGAGCAACGAT 1867
1561 CGTTTGGCGGCGCCCTACGTAACAGATGGGGTGAATGAGACGAGCGTCTCTCAAC 1620
1868 CGTTTGGCGGCGCCCTACGTAACAGATGGGGTGAATGAGACGAGCGTCTCTCAAC 1927
1621 AACACGCGGCGCCACAGGGGCAACTGTTCGCTGTATCATGATGATGAATGACCGGGTTC 1680
1928 AACACGCGGCGCCACAGGGGCAACTGTTCGCTGTATCATGATGATGAATGACCGGGTTC 1987
1681 ACCAAGAGTGTGGGGGCGCCCGCTGCAACATCGGGGGGTTGGGCAACAACATTTGATC 1740
1988 ACCAAGAGTGTGGGGGCGCCCGCTGCAACATCGGGGGGTTGGGCAACAACATTTGATC 2047
1741 TGCCCCACGAGCTGTTCCGGAAGATCCCGAGGCACTTACACCAATGCGGTTGGGG 1800
2048 TGCCCCACGAGCTGTTCCGGAAGATCCCGAGGCACTTACACCAATGCGGTTGGGG 2107
1801 CTTGGTTG 1809
2108 CTTGGTTG 2116

RESULT 2
A32202 2116 bp DNA linear PAT 10-DEC-1996
LOCUS NANBHV PT polypeptide fragment.
DEFINITION A32202
ACCESSION A32202 GI:1926562
VERSION A32202.1
KEYWORDS
SOURCE
ORGANISM Non-A, non-B hepatitis virus
Non-A, non-B hepatitis virus
Viruses; unclassified viruses.
REFERENCE
1 (bases 1 to 2116)
AUTHORS
JOURNAL Patent: FR 2655990-A 21 JUN-1991;
FEATURES
source
1..2116
/organism="Non-A, non-B hepatitis virus"
/mol_type="unassigned DNA"
/db_xref="taxon:12440"
308..>2116
/codon_start=1
CDS

```
/product="partial_NANBH-Pr polypeptide"
/protein_id="CAA02094.1"
/db_xref="GI:1926563"
/translation="MSTNPKPQKTKRNTNRPRQVDFKFGGQIVGGVYLPRRGRL
GVATRTKTSERSPGRROPITKARQPERAWAOPGYPMLVGNEMGMAVLLSPRG
SPSPMGPTDPRRRSRRLKVIDTLTGFDLMGYILPVAGAPLGGABALAHGVRTH
GVNVAITGNLPGCSFSILALISCLTIPASAYEVNRVSGIHYVNDSSNSIYVEDAD
MIMHTPGCVPCYREGNSRRCWALPTLTAAPASITPATIRRHVLDLVAAAFCSMY
VDDCGSVFLVSQLEFSPRRHQTVDDCSSTYPGHVSQHRVMDMMNMSPTALV
SGLRLIQAIVDMVAAGHVAAGLAVYMNVAWVAVLMLEFACVDEPVTGGTH
GHAAGLITSLFTPGPAOKIQLVNTGSMHINRTALNCNLSLGTFLAALFYTHFVAS
GSEBMAVSRPIIDOPQMGPIITYNESHLDORPYCMHAPADPGCVLPALQVCPAYC
FSPSPVVGSTTRFAPATYRKMBENETDVLINNTPRPRGNMFGCTMMNSTGTFTKCGG
PCNIGGVANTLICTPTDPRKHPATYTKCSGSPVL"
```

ORIGIN

```
Query Match      100.0%; Score 1809; DB 6; Length 2116;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1809; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGAGCAGAACTCTTAACCTCAAGAAAAACCAAGTACACCAACGCGCCCAAG 60
308 ATGAGCAGAACTCTTAACCTCAAGAAAAACCAAGTACACCAACGCGCCCAAG 367
61 GAGTCAAGTCCCGGCGGTGTGATCTGTGTGAGTTTAACTGTTCGCGCAG 120
368 GAGTCAAGTCCCGGCGGTGTGATCTGTGTGAGTTTAACTGTTCGCGCAG 427
121 GAGCAGAGTGGGTGTGCGCGAGTAAAGACTTCCAGCGGTGCAACTCGTGA 180
428 GAGCAGAGTGGGTGTGCGCGAGTAAAGACTTCCAGCGGTGCAACTCGTGA 487
181 AGGCGCAACTATCCCAAGGCTGCCAGCGCGAGGCGCTGGGCTCAGCCCGG 240
488 AGGCGCAACTATCCCAAGGCTGCCAGCGCGAGGCGCTGGGCTCAGCCCGG 547
241 TACCTTGGCCCTCTATAGGCAAGAGGATAGGGGTGGGAGAGATGGCTCTGTCA 300
548 TACCTTGGCCCTCTATAGGCAAGAGGATAGGGGTGGGAGAGATGGCTCTGTCA 607
301 CGTGGTCCCGGCTAGTGGGAGCCCACTGACCCCGGCGTGTGAGTTCGCTTA 360
608 CGTGGTCCCGGCTAGTGGGAGCCCACTGACCCCGGCGTGTGAGTTCGCTTA 667
361 AAAGTCATGATACCTCTCAATGCGGCTTCCGCACTCAATGGGATTAATCCGCT 420
668 AAAGTCATGATACCTCTCAATGCGGCTTCCGCACTCAATGGGATTAATCCGCT 727
421 GGGCGTCTCTTAGGGGGCGTGCAGAGGCTTGGCGCATGGCGTCCGGGTTCTGAG 480
728 GGGCGTCTCTTAGGGGGCGTGCAGAGGCTTGGCGCATGGCGTCCGGGTTCTGAG 787
481 GGGCGTCTCTTAGGGGGCGTGCAGAGGCTTGGCGCATGGCGTCCGGGTTCTGAG 540
788 GGGCGTCTCTTAGGGGGCGTGCAGAGGCTTGGCGCATGGCGTCCGGGTTCTGAG 847
541 TTGCTGTCTGTGTCATTCAGCTTCGCTTATGAATGGCGCAACGTGTCGGGAT 600
848 TTGCTGTCTGTGTCATTCAGCTTCGCTTATGAATGGCGCAACGTGTCGGGAT 907
601 TACCATGTCAAGAACTGCTTCAACTCAAGATGTGTACAGACAGGGAGCATGATC 660
908 TACCATGTCAAGAACTGCTTCAACTCAAGATGTGTACAGACAGGGAGCATGATC 967
661 ATGCAACCCCGGGGTGTGAGCTGTGTCGGGAGGGATTAATCCCGGCTGCGGGAT 720
968 ATGCAACCCCGGGGTGTGAGCTGTGTCGGGAGGGATTAATCCCGGCTGCGGGAT 1027
721 GGGCTCACTCCCAAGCTCGCGGCAAGAGCGCCAGATCCCACTGCGCAATATCAAG 780
1028 GGGCTCACTCCCAAGCTCGCGGCAAGAGCGCCAGATCCCACTGCGCAATATCAAG 1087
781 CAGCTGATTTGCTCGTTGGGGCGGCTCCTTCTGCTCCGCTATGTACGTGGGGATCTC 840
```

```
|||||
1088 CAGCTGATTTGCTCGTTGGGGCGGCTGCTCTGTGCTCCGCTATGTACGTGGGGATCTC 1147
841 TGGGATCTGTTTTCTCGTCTCTCAAGCTGTTCACTTCTCGCTCCGCAATCAAGC 900
1148 TCGGATCTGTTTTCTCGTCTCTCAAGCTGTTCACTTCTCGCTCCGCAATCAAGC 1207
901 GTACAGATCTGAATTTGTTCAATCTATCCCGGCAAGTATCAAGTACCGGATGGTTGG 960
1208 GTACAGATCTGAATTTGTTCAATCTATCCCGGCAAGTATCAAGTACCGGATGGTTGG 1267
961 GATATGATGATGAACCTGATCACTCAAGAGCGCTAGTGGTATCCGAGCTACTCCGGATC 1020
1268 GATATGATGATGAACCTGATCACTCAAGAGCGCTAGTGGTATCCGAGCTACTCCGGATC 1327
1021 CCACAAAGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
1328 CCACAAAGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1387
1081 TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
1388 TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1447
1141 GTTGAAGGGGAACCTTACAGACAGAGGGGACACAGCGCGCGCCAGCGGCTTTACA 1200
1448 GTTGAAGGGGAACCTTACAGACAGAGGGGACACAGCGCGCGCCAGCGGCTTTACA 1507
1201 TCCCTCTTCAACCTTGGGCGGCTCGAATAATCCAGCTTTGTAACCAACGCGCAGCTGG 1260
1508 TCCCTCTTCAACCTTGGGCGGCTCGAATAATCCAGCTTTGTAACCAACGCGCAGCTGG 1567
1261 CACATTAACAGAACTGCTTGAATCTGATGATGATGATGATGATGATGATGATGATGATGAT 1320
1568 CACATTAACAGAACTGCTTGAATCTGATGATGATGATGATGATGATGATGATGATGATGAT 1627
1321 CTGTTCTACAGCAAGGTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
1628 CTGTTCTACAGCAAGGTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1687
1381 CCCATTGACAGATTCGATCAGAGGGGTGGGTGCCATCACTTATATAGTCCCAAGCTTGG 1440
1688 CCCATTGACAGATTCGATCAGAGGGGTGGGTGCCATCACTTATATAGTCCCAAGCTTGG 1747
1441 GACACAGAGGCTCTTATGCTGCACTACGCACTCAACCGTGTGTATGCTGCCCGCTTGG 1500
1748 GACACAGAGGCTCTTATGCTGCACTACGCACTCAACCGTGTGTATGCTGCCCGCTTGG 1807
1501 CAGGTGTGTGCGCCAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
1808 CAGGTGTGTGCGCCAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1867
1561 CGTTTTCGCGCCCTTACGATCAGATGAGGTGAGAAATGACAGGACGTGCTTCTCAAC 1620
1868 CGTTTTCGCGCCCTTACGATCAGATGAGGTGAGAAATGACAGGACGTGCTTCTCAAC 1927
1621 AACACGCGGCTCCGACAGGGGCAACTGCTTCCGCTGTACATGATGATATGACACCGGCTTC 1680
1928 AACACGCGGCTCCGACAGGGGCAACTGCTTCCGCTGTACATGATGATATGACACCGGCTTC 1987
1681 ACCAAGACGTGTGGGGGCGCCCGCGACACATCGGGGGGTGGGCAACCACTTTGATC 1740
1988 ACCAAGACGTGTGGGGGCGCCCGCGACACATCGGGGGGTGGGCAACCACTTTGATC 2047
1741 TGGCCCAAGGATGCTTCCGGAAGATCCCGAGGCACTTACCAACCAATGCGGTTCCGGG 1800
2048 TGGCCCAAGGATGCTTCCGGAAGATCCCGAGGCACTTACCAACCAATGCGGTTCCGGG 2107
1801 CCTTGGTTG 1809
2108 CCTTGGTTG 2116
```

RESULT 3

ARI44050
LOCUS ARI44050 2116 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 21 from patent US 6210675.
ACCESSION ARI44050
VERSION ARI44050.1 GI:15105917
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2116)
AUTHORS Highfield,P.Edmund., Rodgers,B.Colin., Tedder,R.Seton. and
Barbara,J.Anthony.James.
TITLE PT-NANB hepatitis polyepitides
JOURNAL Patent: US 6210675-A 21 03-APR-2001;
FEATURES
source 1.. 2116
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 1809; DB 6; Length 2116;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1809; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGACGCAATCTCTAAACCTCAAGAAAACCAACGTAACCAACCGCCGCCACAG 60
DB 308 ATGACGCAATCTCTAAACCTCAAGAAAACCAACGTAACCAACCGCCGCCACAG 367
QY 61 GACGTCAGTTCGCCGGCGGTGTGTCAGATCGTTGGTGTGAGTTTACTCTGTGCCGGCAG 120
DB 368 GACGTCAGTTCGCCGGCGGTGTGTCAGATCGTTGGTGTGAGTTTACTCTGTGCCGGCAG 427
QY 121 GACCCAGGTTGGGTGTGCGCGCAGACTAGAGAAGACTTCGAGCGGTGCGAACCTGTGGA 180
DB 428 GACCCAGGTTGGGTGTGCGCGCAGACTAGAGAAGACTTCGAGCGGTGCGAACCTGTGGA 487
QY 181 AGCGCAACAACCTATCTCCCAAGGCTGCGCAGCCGAGGGGCGCTGGGCTCAGCCCGG 240
DB 488 AGCGCAACAACCTATCTCCCAAGGCTGCGCAGCCGAGGGGCGCTGGGCTCAGCCCGG 547
QY 241 TACCCCTGGCCCTCTATGCGCAACGAGGCGATGGGTGGGCAAGATGGCTCTGTCAACC 300
DB 548 TACCCCTGGCCCTCTATGCGCAACGAGGCGATGGGTGGGCAAGATGGCTCTGTCAACC 607
QY 301 CGTGGCTCCGGGCTAGTGGGGCCCACTGACCCCGGCGTAGGTCGCGTAATTGGGT 360
DB 608 CGTGGCTCCGGGCTAGTGGGGCCCACTGACCCCGGCGTAGGTCGCGTAATTGGGT 667
QY 361 AAAGTCATGATTAACCTCAATGCGGCTTGCCTGACCTTCATGAGGATTAATTCGCTGTC 420
DB 668 AAAGTCATGATTAACCTCAATGCGGCTTGCCTGACCTTCATGAGGATTAATTCGCTGTC 727
QY 421 GCGGCTCCCTTATGAGGGGCGCTGCGAGGCGCTGGGGCGATGGGTCCGGGTTCTGAGAGAC 480
DB 728 GCGGCTCCCTTATGAGGGGCGCTGCGAGGCGCTGGGGCGATGGGTCCGGGTTCTGAGAGAC 787
QY 481 GCGGCTGAACTATGCAACAGGGAATTTAACCGGTTGCTCTTCTCATCTCTCTTGGGT 540
DB 788 GCGGCTGAACTATGCAACAGGGAATTTAACCGGTTGCTCTTCTCATCTCTCTTGGGT 847
QY 541 TTGCTGCTCTGTTTATGCAATTCACGCTTCGCTTATGAAATGCGCAAGTGTCCGGGATC 600
DB 848 TTGCTGCTCTGTTTATGCAATTCACGCTTCGCTTATGAAATGCGCAAGTGTCCGGGATC 907
QY 601 TACCATGTCAGGAAGATTGCTCCAACTTAAGCATGCTGTATGAGAACAGCGGACATGATC 660
DB 908 TACCATGTCAGGAAGATTGCTCCAACTTAAGCATGCTGTATGAGAACAGCGGACATGATC 967
QY 661 ATGCAACACCCCGGGGTGTGCTGTGTCCGGAGGGGTAATTCCTCCCGGCTGTGGGTA 720
DB 968 ATGCAACACCCCGGGGTGTGCTGTGTCCGGAGGGGTAATTCCTCCCGGCTGTGGGTA 1027
QY 721 GCGCTCACTCCACGCTGCGGGCAAGAGCGCAAGATCCCACTGCGCAATACAGACGC 780

DB 1028 GCGCTCACTCCACGCTGCGGGCAAGAGCGCAAGCATCCCACTGCGAATAAGACGC 1087
QY 781 CACGTCGATTTTGTCTGTTGGGGCGGCTTGTCTGCTCCGTATGTAGTGGGGATCTTC 840
DB 1088 CACGTCGATTTTGTCTGTTGGGGCGGCTTGTCTGCTCCGTATGTAGTGGGGATCTTC 1147
QY 841 TCGGATCTGTTTCTCTGCTCTCAAGCTGTTTCACTCTCGGCTCGCGCATGACAG 900
DB 1148 TCGGATCTGTTTCTCTGCTCTCAAGCTGTTTCACTCTCGGCTCGCGCATGACAG 1207
QY 901 GTACAGACTGCAATTTGTTCAATCTATCCGGGCAAGTATCAGGTACCGCATGACTTGG 960
DB 1208 GTACAGACTGCAATTTGTTCAATCTATCCGGGCAAGTATCAGGTACCGCATGACTTGG 1267
QY 961 GATATGATGATGAATCTGTGCACTCAACAGACCCCTAGTGGTATGCGACCTACCGGATC 1020
DB 1268 GATATGATGATGAATCTGTGCACTCAACAGACCCCTAGTGGTATGCGACCTACCGGATC 1327
QY 1021 CCACAGCTGTCTGAGCATGATGGCGGGGCGCACTGGGAGTCTTGGCGGCTTGGC 1080
DB 1328 CCACAGCTGTCTGAGCATGATGGCGGGGCGCACTGGGAGTCTTGGCGGCTTGGC 1387
QY 1081 TACTATTCATGTGGGGAATGGGCTTAAAGGTTTGTGTGATGCTACTCTTGGCCGGC 1140
DB 1388 TACTATTCATGTGGGGAATGGGCTTAAAGGTTTGTGTGATGCTACTCTTGGCCGGC 1447
QY 1141 GTTGAACGGGGAACCTTAAACAGACAGGGGGGACACACGGCGCGGCCCAAGGGCTTACA 1200
DB 1448 GTTGAACGGGGAACCTTAAACAGACAGGGGGGACACACGGCGCGGCCCAAGGGCTTACA 1507
QY 1201 TCCCTCTTCAACACCTGGGCGGCTCAGAAAATTCAGAGTTGTAAACACCAACGAGAGCTGG 1260
DB 1508 TCCCTCTTCAACACCTGGGCGGCTCAGAAAATTCAGAGTTGTAAACACCAACGAGAGCTGG 1567
QY 1261 CACATCAACAGAACTGCTTGAATGCAATGATCTCTCAAACTGGGTTCTTGGCCGG 1320
DB 1568 CACATCAACAGAACTGCTTGAATGCAATGATCTCTCAAACTGGGTTCTTGGCCGG 1627
QY 1321 CTGTTCTTACAGGCAAGGTTCAATGCGGCGGAGTCTCAGAGCGCATGGCCAGCGCCGC 1380
DB 1628 CTGTTCTTACAGGCAAGGTTCAATGCGGCGGAGTCTCAGAGCGCATGGCCAGCGCCGC 1687
QY 1381 CCCATTGACAGTTCGATCAAGAGGTGGGGTCCCATCACTTATTAAGATGCCACGGCTTG 1440
DB 1688 CCCATTGACAGTTCGATCAAGAGGTGGGGTCCCATCACTTATTAAGATGCCACGGCTTG 1747
QY 1441 GACCAAGGCGCTTATTTGCTGCACTACGACCTCAACCGTGTGATCTGTCGGGCTTG 1500
DB 1748 GACCAAGGCGCTTATTTGCTGCACTACGACCTCAACCGTGTGATCTGTCGGGCTTG 1807
QY 1501 CAGGTGTGTGCGCAAGTGTACTGTTTCACTCAAGCCCTGTGTGTGGGGACAGACGAT 1560
DB 1808 CAGGTGTGTGCGCAAGTGTACTGTTTCACTCAAGCCCTGTGTGTGGGGACAGACGAT 1867
QY 1561 CGTTTCGGCGCGCTTACGTAAGATGGGGTGAAGATGAAGACGACGTCTCTCAAC 1620
DB 1868 CGTTTCGGCGCGCTTACGTAAGATGGGGTGAAGATGAAGACGACGTCTCTCAAC 1927
QY 1621 AACACGCGCGCGCAACGGGGCAACTGTTGGCTGTATCATGATGAATAGACCGGGTTC 1680
DB 1928 AACACGCGCGCGCGCAACGGGGCAACTGTTGGCTGTATCATGATGAATAGACCGGGTTC 1987
QY 1681 ACCAAGAGTGTGGGGGCGCCCGCTGCAACATCGGGGGGGTGGCAACAACTTTGATC 1740
DB 1988 ACCAAGAGTGTGGGGGCGCCCGCTGCAACATCGGGGGGGTGGCAACAACTTTGATC 2047
QY 1741 TGCCCCACGAGCTCTTCCGGAAGCATCCGAGGCGCACTTACACCAATGCGGTTCCGGG 1800
DB 2048 TGCCCCACGAGCTCTTCCGGAAGCATCCGAGGCGCACTTACACCAATGCGGTTCCGGG 2107
QY 1801 CTTGGTTG 1809

Db 2108 CTTGTTG 2116

RESULT 4
LOCUS HPCKR12
DEFINITION Hepatitis C virus (strain HCV-1b, clone HCV-K1-R2), complete genome
sequence.

ACCESSION D50481.1
VERSION D50481.1
KEYWORDS GI:1030705
polyprotein; core protein; E1; E2; NS2; NS3; NS4A; NS4B; NS5A;
NS5B; envelope protein; non-structural protein;
interferon-sensitive; interferon-resistant; IFN-sensitive;
IFN-resistant; ISDR; interferon sensitivity determining region;
HVR; hypervariable region.

SOURCE
ORGANISM Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepadnavirinae.

REFERENCE 1 (bases)
AUTHORS Enomoto, N., Sakuma, I., Asahina, Y., Kurosaki, M., Murakami, T., Yamamoto, C., Izumi, N., Marumo, F., and Sato, C.
TITLE Comparison of full-length sequences of interferon-sensitive and resistant hepatitis C virus 1b. Sensitivity to interferon is conferred by amino acid substitutions in the NS5A region
JOURNAL J. Clin. Invest. 96 (1), 224-230 (1995)
MEDLINE 95340824
PUBMED 7542279

REFERENCE 2 (bases 1 to 9410)
AUTHORS Enomoto, N.
JOURNAL Unpublished
AUTHORS 3 (bases 1 to 9410)
TITLE Enomoto, N.
JOURNAL Direct Submision
REFERENCE Submitted (08-MAY-1995) Nobuyuki Enomoto, Tokyo Medical and Dental University, Second Department of Internal Medicine, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113, Japan (E-mail: PXM045220niftyserve.or.jp, Tel:03-3813-6111(ex.3224), Fax:03-3818-7177)

FEATURES
source
location/Qualifiers
1..9410
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/strain="HCV-1b"
/db_xref="taxon:11103"
/clone="HCV-K1-R2"
/note="IFN-resistant HCV from patient 2"
330..9362
/codon_start=1
/product="polyprotein"
/protein_id="BA09072.1"
/db_xref="GI:1814086"
/translation="MSTNPKPQKTKRNTNRBQVYKFGGQIVGVVILPRRGRPL
GYRATKTSERSQPRGROPPIKARQPERAMQPOYPMPLVYNEBGMAGMILSRG
SRPMSGPTDPRRSRLKVIDTLTCGFDLMGILVGVAPLGASRALAHGVRVED
GVNATGNTPGCSFSEILALISCLTIPASVVRASGVYHNTDNCNASTVYEAD
MIMHTPGCTPCREANSSRCWALTPTLAARSSVPTTTRRHVDLLVGAAPCSAMT
VVDLGSVFLVSQLTFSPRRHETVODCNCSYPGVSGSRAMDMWMSPTTALV
SOLRLPOAVMMVMGAGHVGALGLAYVMGMAKVLVIMLLFACVDNTVYSGAG
SHHTQGLVSLFPAQAQOKIQLVNTGSMHINRTALCNDLSLTGFLAFLYTHRFS
GCEPRTASCRSIDAFQNGMPTIYAEFGSDORPCMHYAPRCGIVPASEVCGPYC
FTSPVPVGTTRSGAPYSWGENETDVLINNTTRAPOWNRCTMNGCTMGTCCG
PPCNIGGVGNNTLTCTDPRKRPATYTKCSGPMULTRCWVDYIRLHMPTCTNF
TTFKRMVGVGVBHRLNAACNMTGRGCEDEDRSELSPLLSTTEWOLIPCSFTTL
PLVSTGLIHLHONIVDVOYLGVGSAVSVFVIMKWEYILLLFLDLARVACACIMLL
IAOEAALLENLVNLAASVAGAGFLSPVFFCAWYIKGKLVPGAAVFAVYGMPLL
LLALPBPAYAMDREMAASCGAVFVGLLTLSPHYKVLARLIMLLQYFTRABAC
LQVWIPPLNVRGRDAIILLTALHPELISITKILVAILIGPLMVIOAGITMPPYVR
AGLIRACMLVRKVAGHVVQMAFMKALTLTYVVDHITPLDVAHTEGRLDAVAE
PVPFSDMETKITWGAADTAACDITLGLVASARGKEITFGRADSLSGSGMRLAFT
AVSQQTRGLGCTITSLTGRDNQVGEVQVSTAGOSFLATCTNVCMTVYVAGASK
TLAGPKGPTITOMYNTVDOLVWQAPSGARSLTPTCGSDILVTRHADVLPVRRG
DRSGILSPVSYLKSGSGPLCPSGAGVGFRAAVCTCRGAKAVDFTIPESMETT
MSPVPTDSSPPAVPQTFQVHLHAPTSQSKSNKVPVEVAAQGVKVLVINSVAATL

GRGAVSKAHGVDPNIRNGVTRITTTGAPITTSYTGKFLADGGCGSGAVDIIICDECHS
TDSSTIGTGVYLDQDSETRGAVLVUATATPPGSGVTPHPNIEVALNSTGDIPIPKK
AIPIETIKGGRHLIRHSHKKKCDDELAKLSGGINAVATYRGLDVSIVTASGDIYVVA
TALMTGTFGDFDSVIDCNTCVQIVDFSLDPTTLETTVPDADRSORGRGTVVVA
RAGIYRFTVTPGSPGMDPSVLCBQYDGCAMWELTAPETSVRLARAVANTPGLPMWQ
DHLEWESVFTGLTHIDAFLSQTKQAGNPFYLVAQVTCARAAQAPSPDQPMWK
LIRKPTLHGPPTPLIRLGAONENYILTHPIRTKYIMACGADLEVTSTWVVGVLTA
ALAACTGTGSVIVIGRIITLISGKPAIIPDRVLYEEDMEGASLPIYIISGMOLAE
QFKOKALGILOTTATKOAEAAAVVSKMOALTTPAKHMMNTISGQIYAGSTLDEGN
PAIASIMAFATASITSPLTQHTLIRNLIISLGGWAQADLAPRAAFVAGAIAGAVSAI
GLKYLVIDIAGVAGVAGALVAFVMSGDMSTEDLVNLLPAIISPGALVVGWVQV
ILRHVYGPGEAGVOMNRLIAPASGNHVSPHYVESPDAARVQIISLITDOLK
RHOWNNECSTPGSGMWLRDWDNICVLSDFKTMOSRVLPRLVGELSGORGX
GVWRDGMQTTCPGCAOITGHNKSNRIVGPKCSNMTPTNATYPTGPTSP
APNYSBALMRVAEEVEYRVDGHDHYTGMTDNLKCCQVPAPEFFKELGVRILHR
VAPASPLLRDVEVTFQVGLNVVGSQLECEPEPVAVLTMIDPSHTAETAKRL
ARGSPSLSSASASASLAKTCATCTTHSDSDADILEANLIMROEMGNITRVSEN
KVILDSPEPLRAEEDEREVSIPAILRKSKEPPAMPIMAPDYNPPLIEGMDPY
VPPVHGCEPLPTKAPTIPEPRKRRTVLTSTESALAEIATKTEGSESSAASGT
ATAPDPQSDSDAGSDVESYSMPLEBEPDPPDLSGSWSTVEBAGEEDVYCCMS
YVWGTALITPCAESKLEPINLSLSLHHNMVATITRSASOKOKATPRLQVLD
DHYRVLKEMKAKASTVAKRLISVEACKLTPEPHARSFKGADVRNLSRANLHI
RSWMDLOEDTETPIDITIMAKNEVFCVQPEKGGKPAIIVPDLGYVACEKMALYD
VSTLPOAVMGASVGYOYSPGQRFVFLVAMSKKCPMGFAVDFCTDSTENDIRV
ESISYOCODLAPEARQAIRSLTERLYIGSLPNSGONGVRCRCSAGVLTTCGNTL
TCTYKATACRAKAKIQCCTMLVCGDDLVYICSACTQEBAAIARFTEMTKYSAPRG
DPPQPEYDELITLSCSNVSAHDASGKRVYLTDPPTPLKRAAMETLKHPTVMSML
GNIMVAPTLMARMTLMTFHSILAOEOLKALCOIYAGCVSIEPDLPOIOLHL
GLASFHSISYGEINRVASCLRKAFVPLRVMRHRNSVRKALISOGGARAICGVYL
FMVAVRTKLKLTPIPAASQDLSGWFVAGYSGDIIYHSVRARPRFMWCLLISVGV
GYLPLNRR

mat_peptide 330..902
/product="core protein"
mat_peptide 903..1478
/product="E1"
mat_peptide 1479..2756
/product="E2"
mat_peptide 2757..3407
/product="NS2"
mat_peptide 3408..5300
/product="NS3"
mat_peptide 5301..5462
/product="NS4A"
mat_peptide 5463..6245
/product="NS4B"
mat_peptide 6246..7586
/product="NS5A"
mat_peptide 7587..9359
/product="NS5B"
misc_feature 1479..1559
misc_feature 6954..7073
misc_feature /note="interferon sensitivity determining region (ISDR)"

ORIGIN
Query Match 88.8%; Score 1605.8; DB 14; Length 9410;
Best Local Similarity 93.0%; Pred. No. 0;
Matches 1682; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

1 ATGAGCAGCAATCTCTAAACCTCAAGAAACCAACCAACCAACCAACCGCGCCACAG 60
330 ATGAGCAGCAATCTCTAAACCTCAAGAAACCAACCAACCAACCAACCGCGCCACAG 389
61 GACGTCAGATTCGCGGCGGTGTGTCAGATCGTGTGTGAGATTACCTGTTCGCGCAGG 120
390 GACGTCAGATTCGCGGCGGTGTGTCAGATCGTGTGTGAGATTACCTGTTCGCGCAGG 449
121 GGGCCCAAGTTGGGTGTGCGCCGACTAGAGAGCTTCGACCGGTTCGCAACTGCTGGA 180
450 GGGCCCAAGTTGGGTGTGCGCCGACTAGAGAGCTTCGACCGGTTCGCAACTGCTGGA 509
181 AGGCGCAACCTATCCCAAGGCTGCGCGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240

[illegible]

QY	1321	CTGTTTCAACAGCAACAGTTTCAATGCGTCCGGATGCTCAAGGCCATGACCGACGTCGCCG	1380
Db	1650	CTGTTCTACACACACAGGTTAACTCTTCTGAAATGCCAGGCCGATAGCCAGCTGCCG	1709
QY	1381	CCCATTTGACCAAGTTTGCATCAGGGGTGGGGTCCCATCACTTATTAATGATCCCAAGGCTTG	1440
Db	1710	TCGATCGACGGGTTGCTCAGGGGTGGGGTCCCATCACTTACGCTGACCGGGGACGCTCG	1765
QY	1441	GACCAAGAGCCCTATTGCTGCACTACGCACTTCAACCGTGTGATCTGCCCCGCTTG	1500
Db	1770	GACCAAGAGCCCTATTGCTGCACTACGCACTTCAACCGTGTGATCTGCCCCGCTTG	1829
QY	1501	CAGGTGTGTGGCCCAAGTGTACTGTTTTCATCCAAAGCCCTGTGTGTGGGGAGCAACCGAT	1560
Db	1830	GAGGTGTGTGGTCCAGTGTATTTGCTTCAACCCCAAGCCCGTGTGTGGGGAGCAACCGAT	1889
QY	1561	CGTTTGGCGCGCCCTACGTAACAGATGGGGTGAATGAGACGACGTCGCTTCTCAAC	1620
Db	1890	CGTTTGGCGCGCCCTACGTAACGTCGGGGGAGAAATGACAGCACTGTCGCTCTCAAC	1949
QY	1621	AACACGCGGCGGCCACGCGGCGCACTGTTGCGGCTGTACATGATGAATAGACACCGGTTTC	1680
Db	1950	AACACGCGGCGGCCCTCAAGGCAACTGTTGCGGCTGTACATGATGAATAGCACTGCGTTTC	2009
QY	1681	ACCAAGAGTGTGGGGGGCCCCCGGTGCAATCGGGGGGGTGGGCAACAACTTATATC	1740
Db	2010	ACCAAGAGTGTGGGGGGCCCCCGGTGTACATCGGGGGGGTGGGCAACAACTTATATC	2069
QY	1741	TGCCCAACGACGTGCTTCCGAGACATCCGAGGCCACTTACACCAATGCGTTCGGGGG	1800
Db	2070	TGCCCAACGACGTGCTTCCGAGACACCCGAGGCCACTTACACCAATGCGTTCGGGGG	2129
QY	1801	CCTTGCTTG 1809	
Db	2130	CCTTGCTTG 2138	
RESULT 5			
AFJ13916		9359 bp	mRNA linear
LOCUS	AFJ13916		
DEFINITION	Hepatitis C virus polyprotein mRNA, partial cds.		VRL 01-JAN-2002
ACCESSION	AFJ13916		
VERSION	AFJ13916.1	GI:18027684	
KEYWORDS			
SOURCE			
ORGANISM	Hepatitis C virus		
	Hepatitis C virus		
	viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.		
REFERENCE	1 (bases 1 to 9359)		
AUTHORS	Fanning, I.J., Itakura, J., Nagayama, K. and Enomoto, N.		
TITLE	Characteristics of Hepatitis C viral genome associated with disease progression in a homogeneous patient population		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 9359)		
AUTHORS	Fanning, I.J., Itakura, J., Nagayama, K. and Enomoto, N.		
TITLE	Direct Submission		
JOURNAL	Submitted (17-OCT-2000) Medicine, National University of Ireland, Cork, Hepatitis C Unit, Clinical Sciences Building, Cork University Hospital, Cork, Ireland		
FEATURES	Location/Qualifiers		
source	1..9359		
	/organism="Hepatitis C virus"		
	/mol_type="mRNA"		
	/db_xref="taxon:11103"		
	/note="Isolated from source plasma identified as the single source of the iatrogenic infection of Rhesus negative women in Ireland from May 1977-November 1978; exposure was through HCV 1b contaminated anti-D immunoglobulin"		
	1..329		
	330..>9359		
CDS	/codon_start=1		


```
/product="polypeptide"
/protein_id="A155821.1"
/db_xref="GI:18027685"
/translation="MSTNPKPQKTKRNTNRDPQVFPGGQIVGVYLLPRGRRL
SRVATKTSRSGPRGRQPIPKARRPBGTAQPGYIPMLYENEGMAGMALLSPRG
GSRMPTPRRRSRNIGKVIDLTCGFDLMGVIPLVAGFLGAGALAHGVRVLE
GVNATGNLPGGFSIETLALSLCTIPASIEVRNVSRHVVDLCSNATYSKAEAD
MIMTPGCVPCVRENNSRCVATLPTLAARNSIPTTTIRRVVDLVTGAACSMY
VGDLCGVFLVLSQFTFSRPHETVDDCSITPGHSGRMMWDMWMSPTTALV
SOLLRIPQAVDMVAGWGLAGLAYSAGMAYLLYMLFAVDVETETGGAQ
AHTTRGLATLFSFGPQNIQIINTNGSINRALTNCNLSLNGTALFAFYARFAS
GPERMASCRPIDKFAQMGPIYAKPLPQCYMCHAPQCGIVLSPVSGVPGYC
FPSPVGTGTDRFGVPTYSKMGENTDVLINTRPBGWECTMMPGKPTGCG
PRCDIGVNDPLICTPDRKRPBATTYKCGSPMLTPRCVNDYRRLMHTCTANP
TFRKRVYGVGVRHLMACNMTGERCEDRSELSPLSTEMOVLPGSPFTLL
PALSTGLIHLHRIVDVOYLXIGSIVSVALKWEYVLLFLFLAARVACAMMLL
IYQAEALLENLVNLAASVAGTIGLISFLVFCAMYIKRILPQGAAYFYGVYLL
LIALPRAVANDREMAASCGAVFGLALTLSPHYKFLAILWMLQVFTIRABAG
LQWVLPVNRGRDAIILMCAIHPBLIPATIKILAILGLPLWAOAGITRYPYVR
AHGLIPACVLRVAGHYVQNALMLALITGVYVDHITPLRDMHLEGDRDAVVE
PVSFDMETKITTWGADTAAGDIILGLPVSARKREILGDRDSELEGGMRLAIT
ATAQOTRGLAGCIVISLGRDNQVGEVQVSTATQSFATVNCVCAITVHGACTK
TLAQGPIITOMYTNVDODLVGMALPQASVLPCTCGSDILVTRHADVLPVRAG
DSRGLSPRPVSLKSGSGPLCPGSAVGFRAVCTRGKAVDPEVPSMETT
MSPVETDNSPPAVPOTFOVAHLHAPGSGKSTKVPAAVAGYKVLVLPBVAATL
GGAAYMSKAGVDNRTGVRTTTGAPITTSYTGKFLADGCGSGAYDIIICDEHS
TGSTTLGIVLDQASTAGRLVLAATIPGSAVTPHNEIEVALSINIGETPRYK
ALPIETIKGGRHLIFCHSRKQDELAKLISGLANVAAYRGDAVSLIPSGGVVVA
TALMPTGTPGDPDSVIDCNYQVDFPSLDPFTIETTVPODAVSRSQRRGTGK
RTGIYRFTVGERSGMPSVLCCECYDAGCAYELTPATSTRLAAYNTPLPYCO
DHLEFEGVFTGLTHIDAHFLSOTKQAGNPEYLVAVQATCGARAPAPSMOAMKC
LIRKPTLHGPPTPLIRLGAQNEVILTHPIFTFMACHSADIENVATSTVLYGVILA
ALAAVLTGTSVIVGRILISGKPAIVPREVLYQFDMEBCATLPIEQMOJAE
QERQKMLGLQTAQKQEAAPVESKMALESFMAKHMFIISGQYAGLSTLNG
PAISLMAFLTASITSLPTQHTLFINILGEMVAQALPSSAASFQAGIAGAVSI
GKGVLDVILIAGGAVAGLVAFKVSGEMSTEDLVNLIPLISGLVAVGVCAAL
ILRRHYGREGAVQOMNRLIAPASRGHNSPTHYVESPAALVLIISLITTOILK
RLHOMINDECSPTSCSMLADWDMLCTVLTDPKTLQSKLPLRLGVPFLSGRTK
GAVRGDGIQOTTCPCGAQITGHVKNSSMIIVGRTCSNCHVGFPIINATVGTCTSP
APNYSRALMRVTAEBLEYOVDFHYVGTNDVNCQWPAEPFTEVGOVRLR
YAPCKPRLREVTFOVGNLYVLSGLCEPPEVAVATLSTMLTDSHTIAEAKRL
ARGSPSLASSSASOLASPLKATCTTRSDPADLIEANLILROMGNITRVESEN
KYVILDSFPLRAEBDEREVSAPALIRLSRKRPRAMPIMARDVNPVLEAKMDPY
VPPVNGCELPPTKAPRIPPRPRKRTVLTSTVSALABLATITGSSBSAIVDSGT
ATAPEPDDDDGSDVDESYSMPLEBPGDPLSDSDSSVSEAEADVCCMS
YVTWGLITPSCAESKLPIINALNSLRLHNNMVATITRSYSORQKXTPRLQYLD
DHYRDLKEMKAKASTYKAKLPEBAECULTPHSKSFYGAQKVRNLSAVNHI
LSVMKDLDETEPIDTTIMAKNEVFCVQPEKGRPARLIYPPDGLGVCEMALYD
VSTLPOAVMGSSYGFQVSRGQVFEFLVNAAMKKMKMPPAADVTRFDSVTESDIRV
ESYVOCCLAPBAROIRSLERLYTIGBPLTRSKONGYRRCRASGVLTITSCNTL
TYLKLASAACRAKLDCTMLVCGDDLVVICSAGTXEDASIRVETEMTRSAPEG
DPEKPEYDELITSCSNVSVAHDAGKGVYVLRDPTPLPAAAEATMRHPVMSL
GNIMYAPTLMAAMIMLTHFIFLILAOOLEKALDCOIGACGSIPELPIIOQLH
GSAFSLHSYSPREINRVAACTRKLGVPVLRVHRARSRAKLLSGORRANCGVYL
ENMAVRTXKLTPRIPASQLDLSWFAVAGSGDIYHSLSRAPRPRFMLCLLLSIVG
GYLPLNR"
```

ORIGIN

```
Query Match      88.1%; Score 1593.6; DB 14; Length 9359;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1674; Conservative 0; Mismatches 135; Indels 0; Gaps 0;
```

QY 1 ATGAGCAGCAATCTTAACTCAAGAAAAACCAACGTAACCAACCGCGCCCAAG 60
DB 330 ATGAGCAGCAATCTTAACTCAAGAAAAACCAACGTAACCAACCGCGCCCAAG 389

QY 61 GACGTCAAGTCCCGGCGGTGTGATCGTTGGTGGAGTTAACCGTTGGCGGCGAG 120
DB 390 GACGTCAAGTCCCGGCGGTGTGATCGTTGGTGGAGTTAACCGTTGGCGGCGAG 449

QY 121 GGGCCCAAGTTGGGTGCGCGGCACTAGAGAGACTTCCGAGCGGTGCAACTCGTGA 180
DB 450 GGGCCCAAGTTGGGTGCGCGGCACTAGAGAGACTTCCGAGCGGTGCAACTCGTGA 509

QY 181 AGGCGCAACCTATCCCAAGGTGCGCCAGCCGAGGCGAGGCGCTGAGCCCGGG 240
DB 510 AGGCGCAACCTATCCCAAGGTGCGCCGAGGCGAGGCGAGGCGCTGAGCCCGGG 569

QY 241 TACCTTGGCCCCCTATGAGCAACGAGGCGATGAGGCTGAGGATGAGCTCTGACCC 300
DB 570 TACCTTGGCCCCCTATGAGCAACGAGGCGATGAGGCTGAGGATGAGCTCTGACCC 629

QY 301 CGTGGCTCCCGGCTAGTGGGCGCCCACTGACCCCGGCGTAAAGTCCGCTAATTTGGGT 360
DB 630 CGGCGCTCCCGGCTAGTGGGCGCCCACTGACCCCGGCGTAAAGTCCGCTAATTTGGGT 689

QY 361 AAGTATGATGATACCTTCAATGCGGCTTCCGCGACTCATGAGGATACATTCGCTGCTC 420
DB 690 AAGTATGATGATACCTTCAATGCGGCTTCCGCGACTCATGAGGATACATTCGCTGCTC 749

QY 421 GGGCGTCCCTTAGGGGGGCGCTGCGAGGGGCTTGGCGCATGAGCGTCCGGGTCTGAGAGAC 480
DB 750 GGGCGTCCCTTAGGGGGGCGCTGCGAGGGGCTTGGCGCATGAGCGTCCGGGTCTGAGAGAC 809

QY 481 GGGGTGAATGCAACAGGAAATTTACCAGTGTCTTTCTATCTTCTTGGCT 540
DB 810 GGGGTGAATGCAACAGGAAATTTACCAGTGTCTTTCTATCTTCTTGGCT 869

QY 541 TTGCTGCTGTTTGAACATTCAGCTTCCGCTTATGAGAGTGGCGCAAGTGTCCGGATC 600
DB 870 TTGCTGCTGTTTGAACATTCAGCTTCCGCTTATGAGAGTGGCGCAAGTGTCCGGATC 929

QY 601 TACCATGTCAGCAACGATGCTCAACTCAAGCATGCTGACAGAGCGGACATGATC 660
DB 930 TACCATGTCAGCAACGATGCTCAACTCAAGCATGCTGACAGAGCGGACATGATC 989

QY 661 ATGCAACCCCGGGGTGTGCTGCTGTCGCGAGAGGATTAATCTCCGCTGCTGGTA 720
DB 990 ATGCAACCCCGGGGTGTGCTGCTGTCGCGAGAGGATTAATCTCCGCTGCTGGTA 1049

QY 721 GGGCTACTCCCAAGCTGCGCGGCAAGAGCGGACGATCCCACTGCGCAATACAGCG 780
DB 1050 GGGCTACTCCCAAGCTGCGCGGCAAGAGCGGACGATCCCACTGCGCAATACAGCG 1109

QY 781 CAGGTGATTTGCTGTTGGGGGCGGTGCTTCTGCTCGCTATGATGAGTGGGGGATCTC 840
DB 1110 CAGGTGATTTGCTGTTGGGGGCGGTGCTTCTGCTCGCTATGATGAGTGGGGGATCTC 1169

QY 841 TGGGATCTGTTTCTGCTCTCTCAAGCTTTCGCTTCTGCGCGACATCAGACG 900
DB 1170 TGGGATCTGTTTCTGCTCTCTCAAGCTTTCGCTTCTGCGCGACATCAGACG 1229

QY 901 GTRACGAGTGCATTTGTTCAATCTATCCCGGCAAGTATCAGGTACCGCATGCTTGG 960
DB 1230 GTRACGAGTGCATTTGTTCAATCTATCCCGGCAAGTATCAGGTACCGCATGCTTGG 1289

QY 961 GATATGATGATGAACGTGCTCAGTACAGAGCGCTAGTGTATGCGAGCTACCTCGGATC 1020
DB 1290 GATATGATGATGAACGTGCTCAGTACAGAGCGCTAGTGTATGCGAGCTACCTCGGATC 1349

QY 1021 CCAAGACTGTGCTGAGCAATGCTGCGGGGCGCACTGAGGAGTCTTGGCGGCTTGGC 1080
DB 1350 CCAAGACTGTGCTGAGCAATGCTGCGGGGCGCACTGAGGAGTCTTGGCGGCTTGGC 1409

QY 1081 TACTATTCATGCTGAGGAACTGAGGCTTAAGGCTTGTGATGATGCTACTTTTGGCGGC 1140
DB 1410 TACTATTCATGCTGAGGAACTGAGGCTTGTGATGATGCTACTTTTGGCGGC 1469

QY 1141 GTTGAAGGAGAACTTACAGAGCAGGAGGAGACACAGCGCGCGGCGCCAGCGGCTTAA 1200
DB 1470 GTTGAAGGAGAACTTACAGAGGAGGAGACACAGCGCGGCGCCAGCGGCTTAA 1529

QY 1201 TCCCTTTTCAACTGAGCGGCTCAGAAATCAGCTTGTAAACCAACGCGAGCTGG 1260
DB 1530 ACCCTCTTTTGTGGGGCGTCCCAAAACATCGACTTAACCAACGCGAGCTGG 1589

Qy 1261 CACATCAACAGACTGCTTGAATGCAATGACTCCCTTCAAACTGGGTTCTTGGCCGCG 1320
| | | | |
Db 1590 CACATCAACAGACTGCTTGAATGCAATGACTCCCTTCAAACTGGGTTCTTGGCCGCG 1649
| | | | |
Qy 1321 CTGTTCTACACGCAAGGTTCAATGCTCCGATGCTCAGACGCACTGGCCAGCTCCGC 1380
| | | | |
Db 1650 CTGTTCTACACGCAAGGTTCAATGCTCCGATGCTCAGACGCACTGGCCAGCTCCGC 1709
| | | | |
Qy 1381 CCCATTGACAGTTGCTCAGAGGGGTGGGGTCCCATCTATTATATAGATCCCAAGCGCTTG 1440
| | | | |
Db 1710 CCCATTGACAGTTGCTCAGAGGGGTGGGGTCCCATCTATTATAGATCCCAAGCGCTTG 1769
| | | | |
Qy 1441 GACCAAGAGGCTTATTTCTGGAAGTACGCACTCAACCGTGTGATATCGTCCCGCTTG 1500
| | | | |
Db 1770 GACCAAGAGGCTTATTTCTGGAAGTACGCACTCAACCGTGTGATATCGTCCCGCTTG 1829
| | | | |
Qy 1501 CAGGTGTGAGGCGCCAGTGTACTGTTTCACTCCAGGCGCTTGTGTGTGGGGACGACCAT 1560
| | | | |
Db 1830 CAGGTGTGAGGCGCCAGTGTACTGTTTCACTCCAGGCGCTTGTGTGTGGGGACGACCAT 1889
| | | | |
Qy 1561 CGTTTCGGGCGCCCTACGATGAGATGGGGTGAAGATGAGACGAGCGTGTCTTCTCAAC 1620
| | | | |
Db 1890 CGTTTCGGGCGCCCTACGATGAGATGGGGTGAAGATGAGACGAGCGTGTCTTCTCAAC 1949
| | | | |
Qy 1621 AACACGCGGCGCCCAAGGCGCACTGTTTGGCTGTACATGATGATAGCACTGGGTTTC 1680
| | | | |
Db 1950 AACACGCGGCGCCCAAGGCGCACTGTTTGGCTGTACATGATGATAGCACTGGGTTTC 2009
| | | | |
Qy 1681 ACCAAGAGCTGTGGGGGCGCCCGCTGTCAATTCGGGGGGGTGGGCAACACTTTGATC 1740
| | | | |
Db 2010 ACCAAGAGCTGTGGGGGCGCCCGCTGTCAATTCGGGGGGGTGGGCAACACTTTGATC 2069
| | | | |
Qy 1741 TSCCCCAAGCACTGTTCCGGAAGATCCCGAGGCACTTACACCAATGCGGTTGCGGG 1800
| | | | |
Db 2070 TSCCCCAAGCACTGTTCCGGAAGATCCCGAGGCACTTACACCAATGCGGTTGCGGG 2129
| | | | |
Qy 1801 CCTTGGTTG 1809
| | | | |
Db 2130 CCTTGGTTG 2138
| | | | |
RESULT 6
HPCK152
LOCUS Hepatitis C virus (strain HCV-1b, clone HCV-K1-S2), complete genome
DEFINITION
sequence.
ACCESSION D50485
KEYWORDS
D50485.1 GI:1030704
polypeptide; core protein; E1; E2; NS2; NS3; NS4A; NS4B; NS5A;
NS5B; envelope protein; non-structural protein;
interferon-sensitive; interferon-resistant; IFN-sensitive;
IFN-resistant; ISDR; interferon sensitivity determining region;
HVR; hypervariable region.
SOURCE
Hepatitis C virus
Hepatitis C virus
ORGANISM
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepadnaviruses.
REFERENCE
1 (sites)
Enomoto, N., Sakuma, I., Asahina, Y., Kuroaki, M., Murakami, T.,
Yamamoto, C., Izumi, N., Marumo, F. and Sato, C.,
Comparison of full-length sequences of interferon-sensitive and
resistant hepatitis C virus 1b. Sensitivity to interferon is
conferred by amino acid substitutions in the NS5A region
J. Clin. Invest. 96 (1), 224-230 (1995)
JOURNAL MEDLINE
PUBMED 7542279
2 (bases 1 to 9410)
Enomoto, N.
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 9410)
Enomoto, N.
AUTHORS Direct Submission
TITLE Submitted (08-MAY-1995) Nobuyuki Enomoto, Tokyo Medical and Dental

University, Second Department of Internal Medicine; 1-5-45 Yushima,
Bunkyo-ku, Tokyo 113, Japan (E-mail: ENOMOTO@niftyserve.or.jp,
Tel:03-3813-6111 (ex.3224), Fax:03-3818-1777)

FEATURES

source

1. 9410
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/strain="HCV-1b"
/db_xref="taxon:11103"
/clone="HCV-K1-S2"
/note="IFN-sensitive HCV from patient 2"
330..9362
/codon_start=1
/product="polyprotein"
/protein_id="BA09076.1"
/db_xref="GI:1814089"

CDS

1
/translation="MSTNPKPQKTKNTNRPPDVKPRGGGQIVGGVYLLIPRGRPL
GVRAIRKTSERSQRRGRQPIPKAQPEGRAWAPQYWPVLYNGMGWAGWLLSPRG
SRPSPGTPDPRRRSRNLGKVIDTLCPGADLMGYIPIVGAFLGASRALAGVLEAD
GVNATGNTLPQCSFPIFLALISCLTIPASAYEVNAGSVYHTNDCSNASIVYEAD
MIMHPGCVPCVRANSRCWVALLPILANRSVPTTIRRHVDLLVGAAPGSAVY
VGLDGSVFLVLSOLFTRSPRHETVODCNCSYRGVSHGMDMMDMNSPTALVY
SQLRIPQAVMDVAGAGWGLAATYSWGNMAKYLIVMLPAGVDGRITVTGGAQ
GHTTQRLASLFTGAQQRIDLVNTGSMHINRTALNCNDSLNTGFLAALFYTHFNSS
GCPBRASCRSIDAPQWGPITYAESDSSDORPCWHYAPRPGIIPASBVCPIVC
FTSPVVGTTDRSGAPYSWGENETVLLNTRARQGMFGCTMANGTGTCTCGG
PCNIGVGNNTLTCCPTDCPRKHBEATYTKGSPMWTLPKCMVYIPRLIMYPCVNF
TIPKRYVHGVEHRIANACNMPGRCEDREDRSELSPLISTTQQILPSCFTTL
PVSTGLIHQNIIVKDYVLYGVSAVVSFTIKKAYITLLFLLLADNRACGLMMMLL
IQAEGALENLVNLNLAASVAGAHGFLSELPFCKAWYIKGLVFGAAVAFPGWPLL
LQMLIPRAVYMDREMASSCGGAFVGLALTLTSPHKVYLARLIMLYOFTPEAC
LQWLIPLNLVNGRDAIILTLGALPELISFTKILVALIIGPLMVLQAGITKMPFYR
AOGILRAQMLVRYKAGGHVYOMAKLALITGVYDHLPLQDMAHAGJBLDAVYR
PVSPSMDERTKIITWGADPTAACGDIILGLAASARGKXIFLGPADSLSGGRLLAIPIT
AYSQTTGALACITITSLTGKRXQVBEVQVYSTATSPFLATCTNCTWYTHRAGSK
TLASGKPIITQMYTNVQDILVGMQAPSGASLITCTGSSDLYLVTBHDVTPVRRG
DGRSLSLPRPVSLTKGSSGGPLCPGSHAVGIPRAVCTRGVAKAVDPIPEBEMET
MRSPTVDNNSPPPAVQPTQVYALHAATPSSGSKNKPVEYPAQGYVLPVNSVAATL
GFGAYMSKAGVDNIRGVRTITGAPITGTSTGKFLADGCGGAYDIIICDECS
TDSITLIGITVLDOAETAGARLVLAATPPGVSIVTPHPIEBVALISGEIIPYCK
AIPETIKGGRHLIFGSKKKCBELAKSLGGINAVATRGDVSITPASGDVYVVA
TDALMTGTGTDVSDVLDNCTCVITQVPSLDLPTTIFRTTVPQDAVSRQGRGTGGA
RAGIYRVATPGEPRSGMFDSSVLCBCYDAGCAWELPAETSVRLAYLNTPLQVWQ
DHLEFWSVFGLTHIDAHFLSORQKAGDNPPYVAAQATVCARAAAPRPSMDMMKC
LIRKPLTGHPTPLVLYAGVONVILITHTPIITKIMSGADLEVTSTWLVGCVLA
ALAAVCTLTGSGVIGRIILSGKRALIPDEVLVIREDMESCAHSPYIRGQMOLAR
QPKQKAGLQOTATKQABAAAPVYESKQALETWACGMNFIPTGTOYLAGSLTLPCH
PAISLMAFTASITSPILTQHTLPLNLLGCVAAQLAPPRASAFVGGAGVGA
GLAGVLDVILAGVAGVAGALVAPKWSGMDPSTEDLVNLLPALSFGALVGVCAA
ILRHVHPGEGAVGMNRLIAFASRGHVSFTHVPSPDAARATQIILSLITLQLK
RLHMINDEGSTPCSGSMLRBNVMICTVLDFKTMQSRVLPPLPGVPFLSCGRYK
GVMGDGMQITTCPCGQOITGHVKNKGRMIVGTPRTGCPVAPRPAPEFFKLDGTRLR
APNYSRLMRVAAEYVETRVGDFHYVTGTTDNLCPCQVYAPRPFKELDGRYLR
YAPSKDLRLDEVTFQGLNOYVVSQPCPEBPDVALVLSMLTDPHITLAEAKRLI
ARGSPSILASSAQLSAPSLKATCTTRHSPPDILBANLWQEMGNITRRESN
KVVLDFEPEPRAEEDEREVSLPAEILIRKSRKPPAPRIWAPDYNPLTESMDPY
VPPVHGCPLPTKAPPILPBRKRTVLTFTSSTYSSALAEATPSSSEBSVADSGT
ATAPDQSPNDGAGSDVESWMPLEBGGDDPLSDGSGSTVSESSVACNSMS
YTTTGALITPRAAESKLPINALNSILRHNMYATATSSASGRQKVTPTDRLOVLD
DHYADVLEKMKASITVAKRLVSEACKLPSPASAKSGKARLIVFDLDRVCEKRALDI
RSWKDLIEDETEDITITIAKNENFCVQEBQKSPARLIPGDLDRVCEKRALDI
VSTLPAVWMPSPGYOYSPQORVPLVNMASKKCMGMYDRCDSTVTEGDIIV
EESIYOCDCAPBEARQARSITELVYIGGFLPMNSKNGCGRRARSAGVLTSGCNTL
TCYKATRAACRAATLQDCMTVCGDDILVVICESQOEDBAASLAVPEAMRYVAPSG
DPPOPEYDLEITSCSNVSVAHDASGRKYVYLRDPTPLARAMEYATRPVNSHL
GNITMYATPMAAMIMLTHEPSILLAQQLKALDCQIYGACYSIEPLDQIITQRLH
GLSAFSLHSYSPEINRVAQCLRGLGVPLARVHRARSVAKALISQGRALISGKTL
FNMAVTRTKLTLPIPAASQDLDSGFVAGVSGDIIYHSVSRAPRBMWICLLISVGV
GIVLPPNR"
330..902
/product="core protein"
903..1478
/product="E1"

```
mat_peptide 1479..2756
                /product="E2"
mat_peptide 2757..3407
                /product="NS2"
mat_peptide 3408..5300
                /product="NS3"
mat_peptide 5301..5462
                /product="NS4A"
mat_peptide 5463..6245
                /product="NS4B"
mat_peptide 6246..7586
                /product="NS5A"
mat_peptide 7587..9359
                /product="NS5B"
misc_feature 1479..1559
                /note="Hypervariable region (HVR)"
misc_feature 6954..7073
                /note="Interferon sensitivity determining region (ISDR)"
ORIGIN
```

```
Query Match 88.1%; Score 1593; DB 14; Length 9410;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1674; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

QY 1 ATGAGCAGCAATCTTAACCTCAAGAAAAACCAACGTAACACCAACGCGCCGACAG 60
DB 330 ATAGCAGCAATCTTAACCTCAAGAAAAACCAACGTAACACCAACGCGCCGACAG 389
QY 61 GAGGTCAAGTTCCCGGCGGTGTGTCAGATGTTGTGTGAGTTTACCTGTTGCCGACAG 120
DB 390 GAGGTCAAGTTCCCGGCGGTGTGTCAGATGTTGTGTGAGTTTACCTGTTGCCGACAG 449
QY 121 GGGCCCAAGTTGGGTGTGGCGCGGACTAGAGAACTTCCGAGCGGTGCAACTCGTGA 180
DB 450 GGGCCCAAGTTGGGTGTGGCGCGGACTAGAGAACTTCCGAGCGGTGCAACTCGTGA 509
QY 181 AGGCGACAACCTATCCCGCAAGGCTGCGACGCCGAGGAGCGGCGTGGGCTCAGCCCGG 240
DB 510 AGGCGACAACCTATCCCGCAAGGCTGCGACGCCGAGGAGCGGCGTGGGCTCAGCCCGG 569
QY 241 TACCTTGGCCCTCTATAGCAACGAGGAGTGGGTGGCAGAGATGGCTCTGTGACCC 300
DB 570 TACCTTGGCCCTCTATAGCAACGAGGAGTGGGTGGCAGAGATGGCTCTGTGACCC 629
QY 301 CGTGGCTCCCGGCTTATGTTGGGCCCCCATGACCCCGCGGTAGTGGGTATTTGGT 360
DB 630 CGTGGCTCCCGGCTTATGTTGGGCCCCCATGACCCCGCGGTAGTGGGTATTTGGT 689
QY 361 AAGATGATGATACCTCTCATGAGGCTTGGGAGCTCATGGGGTATCATTCGGCTGTC 420
DB 690 AAGATGATGATACCTCTCATGAGGCTTGGGAGCTCATGGGGTATCATTCGGCTGTC 749
QY 421 GGGCTCCCTTAAGGGGCGCTGCGAGGAGCTGGGCGCATGGCTCGGGTCTGAGAGAC 480
DB 750 GGGCTCCCTTAAGGGGCGCTTCAAGGCGCTGGGCGCATGGTTCGGGTTCTGAGAGAC 809
QY 481 GGGGTAACTAGCAACAGGAAATTTAACCCGTTGCTCTTCTATCTTCTCTTGGCT 540
DB 810 GGGGTAACTAGCAACAGGAAATTTAACCCGTTGCTCTTCTATCTTCTCTTGGCT 869
QY 541 TTGCTGTCTGTTTGAACATTCAGCTTCGCTTAAGTATGGGCAACGTTCCGGGATC 600
DB 870 TTGCTGTCTGTTTGAACATTCAGCTTCGCTTAAGTATGGGCAACGTTCCGGGATC 929
QY 601 TACCATGTACGAAACGATTGCTCAACTCAAGCATGTGTACGAGACAGCGACATGATC 660
DB 930 TATCATGTACGAAACGATGCTTCAAGCAAGCATGTGTATGAGGACAGGACATGATC 989
QY 661 ATGCAACACCCCGGGGTGTGCTGTGTGTCGGGAGGGTATTTCTTCGGTGTGGGT 720
DB 990 ATGCAACACCCCGGGGTGTGCTGTGTGTCGGGAGGGCAATTTCTTCGGTGTGGGT 1049
QY 721 GCGCTACTCTCCAGGCTCGCGGCAAGAGACGACGATCCCACTGGACATATACGACG 780
```

```
DB 1050 GGGCTACTCTCCAGGCTCGCGGCAAGAGACTCCAGGCTCCCACTACGACATATACGACG 1109
QY 781 CAGTCGATTTGCTGTTGGGGGCGCTGCTTCTGCTCCGCTATGTAGTGGGGATCTC 840
DB 1110 CAGTCGATTTGCTGTTGGGGGCGCTGCTTCTGCTCCGCTATGTAGTGGGGATCTC 1169
QY 841 TGGGATCTGTTTCTGCTCTGCTCATGAGTGTACCTTCTGCTGCGGACATCAGAG 900
DB 1170 TGGGATCTGTTTCTGCTCTGCTCTGCTCATGAGTGTACCTTCTGCTGCGGACATCAGAG 1229
QY 901 GTACAGACTGCAATGTTCAATCTATCCCGGCAAGTATCAGAGTACCGGACGCTTGG 960
DB 1230 GTACAGAGTTGCAATGTTCAATCTATCCCGGCAAGTATCAGAGTACCGGACGCTTGG 1289
QY 961 GATATGATGATGAACTGTACCTTACAGAGCCCTATGATGTATCGAGCTACTCCGATC 1020
DB 1290 GATATGATGATGAACTGTACCTTACAGAGCCCTATGATGTATCGAGCTACTCCGATC 1349
QY 1021 CCACAGCTGTGCTGTCACATGATGGCGGGGCGCACTGGGGAGTCTTGGCGGCTTGGC 1080
DB 1350 CCACAGCTGTGTCACATGATGGCGGGGCGCACTGGGGAGTCTTGGCGGCTTGGC 1409
QY 1081 TACTATTTCCATGGTGGGAACTGGGCTAAGTCTTGTGTGTATGCTACTTTTGGCGGC 1140
DB 1410 TACTATTTCCATGGTGGGAACTGGGCTAAGTCTTGTGTGTATGCTACTTTTGGCGGC 1469
QY 1141 GTTGAAGGGAACTTTACAGACAGAGGGGACACAGCGCGCGCCACGAGCTTACA 1200
DB 1470 GTTGAATGGGCGCACACCGTGAACAGGGGGGCGAGGGCCATATACCAACCAAGGCTCGCG 1529
QY 1201 TCCCTTTTCAACACCTGGGCGGCTCAGAAATTCAGCTTGTAAACCAACGCGAGCTGG 1260
DB 1530 TCCCTTTTCAACATTTGGGGCGAGACAGAAATTCAGCTTGTAAATCAACGCGAGCTGG 1589
QY 1261 CACATCAACAGAACTGCTTGAATGCTGAATGCTCCCTCAAACTGGGTTCTTGGCGCG 1320
DB 1590 CACATCAACAGAACTGCTTGAATGCTGAATGCTCCCTCAAACTGGGTTCTTGGCGCG 1649
QY 1321 CTGTTCTACAGCAACGATTCAATGCTCGGATGCTCAGAGCGCATGGCCAGCTGCGC 1380
DB 1650 CTGTTCTACACACAGGTTCAATGCTCGGATGCTCAGAGCGCATGGCCAGCTGCGC 1709
QY 1381 CCCATTGACCACTTGCATAGAGGGGTGGGTCCCATCATCTTATATAGTCCACGCTTG 1440
DB 1710 TCCATCGAGCGTTCGCTCAGGGGTGGGTCCCATCATCTTACGCTGAGCGGGCAGCTGG 1769
QY 1441 GACAGAGGCGCTTATGCTGGCACTACGACCTCAACCGGTGTGATCGTGGCGGCTTG 1500
DB 1770 GACAGAGGCGCTTATGCTGGCACTACGACCTCAACCGGTGTGATCGTGGCGGCTTG 1829
QY 1501 CAGGTGTGTGGCCCAAGTACTGTTTCACTCCAGCCCTGTGTGTGGTGGGACGACGAT 1560
DB 1830 GAGGTGTGTGTGTCAAGTATGTTTCACTCCAGCCCTGTGTGTGGTGGGACGACGAT 1889
QY 1561 CGTTTGGGCGCCCTTACGTAACGATGGGTGGAATGAGACGAGCGTGTCTCTCAAC 1620
DB 1890 CGTTTGGGCGCCCTTACGTAACGATGGGTGGAATGAGACGAGCGTGTCTCTCAAC 1949
QY 1621 AACAGCGGCGCGCAACGCGGCAACTGGTTCGAGCTGTAACATGATGAATGACACCGGGTTC 1680
DB 1950 AACAGCGGCGCGCGCAACGAGGCAACTGGTTCGAGCTGTAACATGATGAATGACACTGGGTTTC 2009
QY 1681 ACCAAGACGTGTGGGGGCGCCCGGTGCAACATCGGGGGGTGGCAACAACCTTTGATC 1740
DB 2010 ACCAAGACGTGTGGGGGCGCCCGGTGTAACATCGGGGGGTGGCAACAACCTTTGATC 2069
QY 1741 TGGCCCAAGGACTGCTTCCGGAAGCATCCGAGGCGCACTTAAACCAATGCGGTTCCGGG 1800
DB 2070 TGGCCCAAGGACTGCTTCCGGAAGCATCCCGAGGCGCACTTAAACCAATGCGGTTCCGGG 2129
QY 1801 CCTTGTTG 1809
```

Db 2130 CCTGCTTG 2138

RESULT 7
LOCUS HCU45476 9431 bp RNA linear VRL 07-FEB-1996
DEFINITION Hepatitis C virus isolate HD-1, complete genome.
ACCESSION U45476
VERSION U45476.1 GI:1181831
KEYWORDS
SOURCE
ORGANISM Hepatitis C virus
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepadnavirinae.
1 (bases 1 to 9431)
Muller.H.M., Pfeiff.B., Heller.A.B., Goesser.T. and Theilmann.L.
Complete sequence of a German isolate of hepatitis C virus (HCV)
derived from a serum of a patient after liver transplantation due
to HCV induced hepatocellular carcinoma
Unpublished
2 (bases 1 to 9431)
Muller.H.M.
JOURNAL Direct Submission
REFERENCE Submitted (11-JAN-1996) Hubert M. Muller, Federal Research Centre
AUTHORS for Virus Diseases of Animals, Paul-Ehrlich Str. 28, P.O. Box 1149,
JOURNAL Tuebingen, D-72001, Germany
TITLE Location/Qualifiers
FEATURES
source 1..9431
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/isolate="HD-1"
/db_xref="taxon:11103"
/note="genotype 1b, isolated from human serum"
1..341
342..9374
/codon_start=1
/product="polyprotein"
/protein_id="AAA86907.1"
/db_xref="GI:1181832"
/translation="MSTNPKPQKTKRNTNRPRDQVKKPGGQIVGVLLPRGRPL
GVRAIRKTSERQSGRRRLIPKAKQPBRSKRAQPGYPMPLIGNMGMAKRLSPRG
SPSWGPTDPRKSRNLKVIDLTGFDLMGYLPVGAIPAGARLALHGVRLRG
GVNVAITGNLPGCSFSIPLIALSLCLTIPASAEVNRNVGVNHTVNDLNSIVETAD
MIMHTPGVPCVREBNSRSCWALPTPLAARNASIPITIRHVLVLGAAGFCAM
VGDLCGSEVLVQLFTFSPRRAVODQCSIPGHSRRAAMNDGMSSTALVAV
SOLRLIPAVVMVAGAHKGVLAGLAAYVMGNMAKVLIMVLLPAGUDCTTVNGSGO
ARTVYELISLITRGPQRILQVNTNGSMHINFTALNCNDSLQETLALPTTRPNS
GCPERMAKSGSIDQFDQMGPIVAPEDLDORPYCMHAPPCGIVPASQVCGPVY
FTSPVVGTTDRSGVPTVSGENETDVLINNTRPQGMNPGCTMNSGTGTCG
PCGIGGVGNNTLICPTDCFRKHPATYTRCGSPWLTPRCAVDYPRMLHPCTVNF
TIFKRMVGVGEHLNACMNTGRGCDLEDRESELPLSLSTTEMQVLCSTPTL
PALSTGLIHLHONVVDOYLVGIGSVVSAIKMEYIILLFILLADAVACIMMMLL
INQARALEENVLVLAASLAGHGLISPLVFCAMYIKGRVLPQGAATLVGVMILL
LLALPPEPAVALDREMAKSGGAVFVGLIITLSPHYAFILRLIMVLIQYFTREAL
LQVMLPELVNRGRDAIILLTCAVHGLIFDITRLLLILGLMLVLAQIRVPEFVR
AQLVLRACMLVRKAVAGHYVQMAFMKLAALSTYVYDHLPLRDRHGLDLAVAE
PVVFSAMEFKLITMGADTPAACGDIILGIPVSRGRREILGPADELEGOMBLATIT
AYAOORGLIGCIITSLGRDKNOGVEGVYSTAFATCNAVGVCMYVYHAGSK
TLAGKGPYTMVTVTDODIVGMAKPRARSLTPTCTGSSSDIYLITRHADYIPARRG
DSKGLSPRPVSYLKSGSGGPLCPSGAAGIIPRAACVTRVAAPVFPVESMETT
MESPVFTDSSSPAVQTFQVAMHAPTSGSTTVPAVAAQGVAKVVLNPSVATL
GGAYSMAKAGIDIPNIRGVRTITGAPITVSTYKFTLADSGSGGAVDIIICDECHS
TDSITLIGIVLDOAETAGALVLAATPPGSAVTPVPSNIEBALSNIGIIPFYK
AIPLENIGGRHLIFCHSRKCKDELAAGLGINAVVYRGDLVSVIPSGDVVVA
TDAMTGYTGDSDVITDCVCTQTVYDSDLPFTIETTVYQDVAVSSQRGRPGRC
RMGIRFPTVGERSGMDSVLCESYDGCAMVELTPREASVRALALNTPGLRCO
DHLFEWSEVFTGLTHIDAFISQTKQAGNPFYVAAYATCARQAPPSDVMWKL
LIRKPTLHGPPLRLYGLAVONEVLTHTPIKTFIMKASALAEVTSVTVYGVGVL
ALAAACLTGCAVIVRILISGRPAIVPEREALVQEDMESCAHLPIBEGMOGAA
QPKKALGILQTAQAEAAAVPESKRALEAPAKMNRNLSGIVLAGISTLPGN
PAIASIMAFATASITSPITQTSLDPIILGAWAQLAPSAASAVAGGIAAAGSTI
GAGKLVLDILAGVAGALVAFKMSSEVSTEDLILNLEPILSLPALVGVCAA
ILRHVGPEGEAVVMNRLIAPAFGNHVSPTHYVBPESDAARVTOIISLITTOLEK

ORIGIN
3'UTR
Query Match 87.4%; Score 1581.8; DB 14; Length 9431;
Beet Local Similarity 92.2%; Pred. No. 0;
Matches 1667; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

1 ATGAGCAGAACTCTAAACCTCAAGAAAACCAAGCTAACCAACCGCGCCACAG 60
|||||
342 ATGAGCAGAACTCTAAACCTCAAGAAAACCAAGCTAACCAACCGCGCCACAG 401
|||||
61 GACGTCAAGTCCCGGGGGGTGTCAGATCGTGGTGAATTACCTGTGCGCGCAGG 120
|||||
402 GACGTCAAGTCCCGGGGGGTGTCAGATCGTGGTGAATTACCTGTGCGCGCAGG 461
|||||
121 GCGCCGAGTGGGTGTCGCGCGCAGCTAGAAACTTCCGAGCGGTGCACTCTGTGGA 180
|||||
462 GCGCCGAGTGGGTGTCGCGCGCAGCTAGAAACTTCCGAGCGGTGCACTCTGTGGA 521
|||||
181 AGGCGCAACCTATCCCGCAAGGCTCGCCAGCGCGAGGCGTGGGCTCAGCGCGGG 240
|||||
522 AGGCGCAACCTATCCCGCAAGGCTCGCCAGCGCGAGGCGTGGGCTCAGCGCGGG 581
|||||
241 TACCTTGGCCCTCTATAGCGCAAGAGGCAATGGAGTGGGAGATGGCTCTGTCAACC 300
|||||
582 TACCTTGGCCCTCTATAGCGCAAGAGGCAATGGAGTGGGAGATGGCTCTGTCAACC 641
|||||
301 CGTGGCTCCCGGCTAGTGGGGGCGCCACTGACCCCGCGGTAGTGCCTGAATTTGGGT 360
|||||
642 CGTGGCTCCCGGCTAGTGGGGGCGCCACTGACCCCGCGGTAGTGCCTGAATTTGGGT 701
|||||
361 AAAGTCATGATACCTCATAGCGGCTTGGCGCACTCATGGGGTATCATTCGCTGCTC 420
|||||
702 AAAGTCATGATACCTCATAGCGGCTTGGCGCACTCATGGGGTATCATTCGCTGCTC 761
|||||
421 GCGGCTCCCTTAGGGGGGCTGCGCAAGGCGCTTGGCGCAATGGCGTTCGGAGAGAC 480
|||||
762 GCGGCGCCCTTAGGGGGGCTGCGCAAGGCGCTTGGCGCAATGGCGTTCGGAGAGAC 821
|||||
481 GCGGCGCACTTAGCGCAAGGAAATTTACCGGCTGCTCTTCTCATATCTTCTGCTGCT 540
|||||
822 GCGGCGCACTTAGCGCAAGGAAATTTACCGGCTGCTCTTCTCATATCTTCTGCTGCT 881
|||||
541 TTGCTGCTCTTGAACCATTCGAGCTTCGCTTATGAAGGCGCAAGTTCGCGGATC 600
|||||
882 TTGCTGCTCTTGAACCATTCGAGCTTCGCTTATGAAGGCGCAAGTTCGCGGATC 941
|||||
601 TACCATGTCAGAAAGATTTGCTCAACTCAAGCATCGTGTACGAGACGCGGACATGATC 660
|||||
942 TACCATGTCAGAAAGATTTGCTCAACTCAAGCATCGTGTGTATGAGACGCGGACATGATC 1001
|||||
661 ATGCAACACCCCGGGGTGTGTCCTGCTGCTGCGGAGGGAATTTCTCTCCGCTGCTGGGTA 720
|||||
1002 ATGCAACACCCCGGGGTGTGTCCTGCTGCTGCGGAGGGAATTTCTCTCCGCTGCTGGGTA 1061
|||||

QY	721	GGGCTCAGCTCCCAAGCTGGGGGCAAGACGCAACATCCCACTGCGCAATATCGAGC	780
Db	1062	GGCTTACCCCAAGCTCGGGCCAGAAACGCAACATCCCTACTACGACATACGAGCT	1121
QY	781	CAGCTGATTTGCTGTTGGGGGAGCTGCTTCTGCTCCGCTATGATGATGGGGATCTC	840
Db	1122	CAGCTGATTTGCTGTTGGGGGAGCTGCTTCTGCTCCGCTATGATGATGGGGATCTC	1181
QY	841	TGGCGATCTGTTTTCTGCTGCTCTTCAGCTGTTCACTTCTGCGCTTCGGCAATCAGAC	900
Db	1182	TGCGGGTCTGTTTTCTGCTGCTCTTCAGCTGTTCACTTCTGCGCTTCGGCAATCAGCA	1241
QY	901	GTAACGAGCTCAATTTGTTCAATCTATCCCGGCCAGTATCAGGTCACCGATGGCTGG	960
Db	1242	GTTACAGAGCTCAATTTGCTCAATCTATCCCGGCCAGTATCAGGTCACCGATGGCTGG	1301
QY	961	GATATGATGATGAATCTGCTCACTTACAGACGCTTATGTTATCGAGCTACTCGGATC	1020
Db	1302	GATATGATGATGAATCTGCTCGCCCAACAGCTTATGTTATCGAGCTACTCGGATC	1361
QY	1021	CCACAGCTGCTGAGACATGATGGCGGGGCCCATGAGGAGTCTGCGGGCTTGGC	1080
Db	1362	CCGCAAGCTGCTGAGACATGATGGCGAGGCGCATGAGGAGTCTGAGCGGCTTGGC	1421
QY	1081	TACTATTTCCATGATGGGGGACCTGGGCTTAAGGCTTGTTGATGCTACTCTTTGCCGCG	1140
Db	1422	TACTATTTCCATGATGGGGGACCTGGGCTTAAGGCTTGTTGATGCTACTCTTTGCCGCG	1481
QY	1141	GTTGACGGGGAACTTTACACGACAGGGGGACACAGCGCGCGGCCCAAGGCTTAC	1200
Db	1482	GTTGACGGGGAACTTTACACGACAGGGGGGACACAGCTCGAACCGTTAAGAACTGACA	1541
QY	1201	TCCCTCTTACACCTGGGGCCGGCTCAGAAAAATCCAGCTTGTAAACCAACGAGAGCTGG	1260
Db	1542	AGCTCTTATACAGCGGGGCCGCTCTCAGAAATCAGCTCTGTAAACCAACGAGAGCTGG	1601
QY	1261	CACATCAACAGAACTGCTTGAATCTGCATGATCTCCCTCCAACTGGGTTCTTTGCCGCG	1320
Db	1602	CACATCAACAGAGCGGCTCTGAATCTGTATGATCTCCCTCCAGACGGGGTTCTTTGCCGCG	1661
QY	1321	CTGTTCTTACAGGCAAGGTTCAATACCGCTCCGAGTCTCAGAGCGGATGCGAGCTGCCG	1380
Db	1662	CTGTTCTTACACACGAGCTTCAACAGCTCTCCGATGCTCCAGAGCGGATGCGAGCTGCCG	1721
QY	1381	CCCATTTGACAGTTTGCATCAGGGGATGGGGTCCCATCACTTAAATGATATCCACAGGCTTG	1440
Db	1722	TCCATTCGACAGTTTGCATCAGGGGATGGGGTCCCATCACTTAAATGATATCCACAGGCTTG	1781
QY	1441	GACCAAGAGGCTTATTTGCTGGCACTACGCACTCAACCTGATCGTATCTGTCCGCGCTTG	1500
Db	1782	GACCAAGAGGCTTATTTGCTGGCACTACGCGCGCACTGTGTGATGTATCCGCGCTTG	1841
QY	1501	CAGGATGTGGGCCAGTGACTGTTTCACTCCAAAGCCCTGTTGATGATGGGAGACGACGAT	1560
Db	1842	CAGGATGTGTGCTCAGGTATTTGCTTCACTCCCAAGCCCTGTTGATGATGGGAGACGACGAT	1901
QY	1561	CGTTTCGGCGCCCTTACGTCACAGATGGGGTGAATGAGACGAGCTGCTCTTCAAC	1620
Db	1902	CGCTTCGGCGCTCCCTTACGTCATAGCTGTGGGGGAGAAATGAGACGAGCTGCTCTTCAAC	1961
QY	1621	AACACGGCGGCGGCAACGGGGGCACTGTGTTCCGCTGTACATGAGATGAATACGACGGGTTT	1680
Db	1962	AACACGGCGGCGGCAACGGGGGCACTGTGTTGCTGTATGATGATGAATACGAGGTTT	2021
QY	1681	ACCAAGAGCTGTGGGGGCCCCCGGTGCAACATCGGGGGGCTCGGCAACAACTTTGATC	1740
Db	2022	ACCAAGAGCTGTGGGGGCCCCCGATGTGAGCATCGGGGGGCTCGGCAACAACTTTGATC	2081
QY	1741	TGCCCCAGGACTGTTCCGAGAGATCCGAGGCACTTACACCAATTCGGTTCCGGG	1800
Db	2082	TGCCCCAGGACTGTTCCGAGAGATCCGAGGCACTTACACCAATGTGGTTCCGGG	2141

Db	1801	CTTGATTG	1809
RESULT 8	AB049095	9585 bp	RNA linear
LOCUS	AB049095	9585 bp	RNA linear
DEFINITION	Hepatitis C virus gene for polypeptide, complete cds.		
ACCESSION	AB049095		
VERSION	AB049095.1	GI:11559456	
KEYWORDS			
SOURCE	Hepatitis C virus		
ORGANISM	Hepatitis C virus		
REFERENCE	1. Hepatitis C virus		
AUTHORS	Takahashi, K., Iwata, K., Matsumoto, M., Matsumoto, H., Nakao, K., Hatahara, T., Ohata, Y., Kanai, K., Maruo, H., Baba, K., Hijikata, M. and Mishiro, S.		
JOURNAL	Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients with hepatocellular carcinoma: the 'progression score' revisited		
REFERENCE	Hepatic Res. 20 (2), 161-171 (2001)		
AUTHORS	2. (bases 1 to 9585)		
JOURNAL	Mishiro, S.		
REFERENCE	Direct Submission		
AUTHORS	Submitted (19-SEP-2000) Shunji Mishiro, Toshiba General Hospital,		
JOURNAL	Department of Medical Sciences, 6-3-22, Higashi Oh-i, Shinagawa-ku,		
REFERENCE	Tokyo 140-8532, Japan (E-mail: shunji.mishiro@po.toshiba.co.jp,		
AUTHORS	Tel: 81-3-3764-8981, Fax: 81-3-3764-8992)		
JOURNAL	Location/Qualifiers		
FEATURES	1. .9585		
source	/organism="Hepatitis C virus"		
	/mol_type="genomic RNA"		
	/isolate="HCV1169"		
	/specific_host="human"		
	/db_xref="taxon:11103"		
	/tissue_type="serum"		
	/note="HCV-genotype: 1b"		
	<1..310		
	/note="partial"		
	311..9343		
	/codon_start=1		
	/product="polypeptide"		
	/protein_id="BA818608.1"		
	/db_xref="GI:11559457"		
	/translation="MSTNPKPQKTKRNTNRDPQVFPQGGQIVGVVLLPRGRPL		
	GVNATRTKTSRSPGRGRQPIPRARPEGAMQOPGMPWLYNBEIGVAGWLLSPG		
	SRSWGPTRRRSRNIGKIDITLTCGDAIMGYIPLVAGYIPLAALAAVGLVLEAD		
	GVNATGNTLPGCSFETLALLSLCLTPASHEVNVGVYTHVINDCSNATVLEAD		
	MINHTPCVPCVRENNSSRCVVALTPTLAARNSSIPATIRHVDILVGAARCSMY		
	VGDTCGVPLVSQLTFSPPRYETVDDNCSTIPGHVSHRNADMMNMSPTTALV		
	SQTLRIEQAVDVNTGAMVGLAYYSVNWGMAVLLVMLFAGVDGTHVYGVS		
	ARNTLTGLTFTSGPAKIQLVNTNGSMFINRALCNLSLNTGFLAALFYTHKFNSS		
	CGSRMASCRIKFGAGWGPITGVSSVDQSPYCMHAYPECGIYPASQVGVYC		
	FTEGSPVVGITDRSGAPYVWGENEDVLLNNTPRPGMGPCGTMMNSIGFPTKGG		
	PCNIGGAVNNTLTCPDRCRKPBEATYTCGSGPMLTPRCIYDIYRIMHYCTANF		
	PIFSVRYVYGVSHRLNAAACNMTGREGCEDLEDRSELSPLLSTTEWQILPESFTLL		
	PAISTGLIHLQNIIVDIQYLYGISAASFAIKMEVYLFEPLLDARVACAMMILL		
	IAQEALENIIVLNLASVAGAGILPFLVFCAMVYIKRLVPAALVAYGVYRMLL		
	LLALPRAVAMDREMAASGAGVFGILILTLSPHYKVLARLLWLOVYFIRAEHL		
	LOWIPIRVNRRGRDAIILLTCMVHSELIPEIKTYILLGLPLVOTGLTRIPVVR		
	AOGLIRICMLVRKAYAGHYQVALLMKLALTYIVYDHLTPRLDMAHAGRIDLAIVE		
	PVPISDMEYKIIITWADTYACGDIILGLPVASARGELILGPADSFGEGWRLLAFTT		
	AYSQQTGILGICITSLTGRDNQVGEVQVSTATQSPFLATCNVGVCTVYHQAQSK		
	TLGAPKPIITOMTYTNDODLVGMPPASGARSLLPTCGSGSDILVLRHADVLEVRRG		
	DSGSLSPRIPIYLGKSSGGPLCPSGHAWGIFRAVCTRGAAKAVDFIPVSMETT		
	MBSGVFLNDSPPAVPOTTPQVALLHAPGSGSKTVPAAYAAAGVYVAVINPEVATL		
	GFQANMSKAGVDPNIRITGRVTTTGAPITVSYGKFLADGCGSGAGYDIITDECHS		
	ITDSITLIGTGVLDQATAGARLVATLPQSVATVPHEINIEVALSTNGELTFPGK		
	AIPETIKGRRLIFGSKKKCKDELLTKVALGVNAVAYYRGIDVSIPTGSGVVVVA		


```

Db      1871 CGTTCCGGCGCCCGCCACGTAACCTTGGGGGAGATGAGACGACCTGCTACTTCTTCAAC 1930
      |||
Qy      1621 AACACCGCGCGCCCGCACGAGGCACTGTTGCGCTGTAACATGATGAAATACACCGGGTTC 1680
      |||
Db      1931 AACACCGCGCGCCCGCACGAGGCACTGTTGCGCTGTAACATGATGAAATACACCGGGTTC 1990
      |||
Qy      1681 ACCAAGAGCTGTCGGGGGCCCCCGCTGCAACATCGGGGGGTTCGGCAACAACACTTTGATC 1740
      |||
Db      1991 ACCAAGAGCTGTCGGGGGCCCCCGCTGCAACATCGGGGGGTTCGGCAACAACACTTTGATC 2050
      |||
Qy      1741 TCGCCCGACGACTGCTTCCGGAAGACATCCCGAGCCCACTTACCAACCAATCGGTTGGGG 1800
      |||
Db      2051 TCGCCCGACGACTGCTTCCGGAAGACATCCCGAGCCCACTTATCAAAATGTTGTTGGGG 2110
      |||
Qy      1801 CCTTGTTG 1809
      |||
Db      2111 CCTTGTTG 2119

RESULT 9
AF207761 9379 bp RNA linear VRL 27-Apr-2000
LOCUS    Hepatitis C virus strain MD20 complete genome.
DEFINITION
ACCESSION AF207761
VERSION   AF207761.1 GI:7650239
KEYWORDS
SOURCE    Hepatitis C virus
          Hepatitis C virus
          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
          Hepacivirus
REFERENCE 1 (bases 1 to 9379)
          Nagayama,K., Kurosaki,M., Enomoto,N., Miyasaka,Y., Marumo,F. and
          Sato,C.
          Characteristics of hepatitis C viral genome associated with disease
          progression
          Unpublished
          2 (bases 1 to 9379)
          Nagayama,K., Kurosaki,M., Enomoto,N., Miyasaka,Y., Izumi,N. and
          Sato,C.
          Direct Submission
          Submitted (23-NOV-1999) Second Department of Internal Medicine,
          Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-ku,
          Tokyo 113-8519, Japan
          Location/Qualifiers
            1. 9379
              /organism="Hepatitis C virus"
              /mol_type="genomic RNA"
              /strain="MD20"
              /db_xref="taxon:11103"
              /note="Isolated from a patient with persistently-normal
              ALT values (ASG-12)"
              ALT values (ASG-12)
              330..9362
              /codon_start=1
              /product="polyprotein"
              /protein_id="AF65951.1"
              /db_xref="GI:7650240"
              /translation="MSTNPKPQRTKRNTRRPODYKFGGGOIVGGVLLPRGRPL
              GYATRTSESRQPRGRROPKPRQPEGMAOPGPMPLINBEGMAGMILSRG
              SRPMGPTDPRRSRLKIVITLTCGPDLMGIVPLVGAALAHGVRVED
              GNVVATGNLPGCSFSLFALLSLCTTPASAVRVNVSQVYLTNCSNASTYEBAD
              MIMHTPGCTPCVRENNSSRCWVALLFTLLARASIPFTTYRRVDDLVGAAACSMY
              VEDLCGSVFLVSQLFTFSRRHETVDCNCSIPGHVTHGRMAWMMNWSPTAALV
              SOLRLPOAVLVNAGAHWGLAGLAYSVNVMWAVLLVMLFAGDGDNTAALV
              AASTTGVSILFSRGANQIKQILINTGSMHINTALNCNBSINTGMLAALYTHKFNAS
              GCEPLRAGSRIIDPOAGMPTHTVTPASDORPCNHYAPRPGCIYPSOVCGPYVC
              FPPSPVVGTTDRFVGTVMGENTDVLILNNSRPLGMPCTMNNSTGPFRTGG
              PCNIGVGNNTITCTPCDCKRKPATYTCGSGPWLTSPCMHYNRILMNYCTNF
              TIKVRIVGGVEHRLNACNMTRGECDEDRSLSPLSLSTTBMOVLPCSTFTL
              PAALSTLILHONIVDVOYLUGISVVVSAIKWEYIILFLILADARVACIMMML
              IADAEALBNLVLINAAVAGAGTSLFVPCAAHYIKRILVPGAAYAGVWPLIL
              LIALPRAVAMREMAAGSCGAVFTGLALTLSPHYKVLARLIMWLYOFSRAEL
              LQWIRPLANVRGRDAIILITCVIHPELITITIKLILAVLGPLMLQAGITKIVFVR

```

ORIGIN

Query Match 87.4%; Score 1580.2; DB 14; Length 9379;
 Best Local Similarity 92.1%; Pred. No. 0;
 Matches 1666; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

```

Qy      1 ATGAGACGAATCTTAACCTCAAGAAAACCAACGTAACCAACCGCGCCGACAG 60
      |||
Db      330 ATGAGACGAATCTTAACCTCAAGAAAACCAACGTAACCAACCGCGCCGACAG 389
      |||
Qy      61 GACGCAAGTCCCGGCGGAGTGCAGATCGTTGAGATTAACTGTTGCCGCGAG 120
      |||
Db      390 GACGCAAGTCCCGGCGGAGTGCAGATCGTTGAGATTAACTGTTGCCGCGAG 449
      |||
Qy      121 GGGCCGAGTTGGGTGGCGCGGAGTAGAAGACTCCGAGCGGTGCAACTGTTGCA 180
      |||
Db      450 GGGCCGAGTTGGGTGGCGCGGAGTAGAAGACTCCGAGCGGTGCAACTGTTGCA 509
      |||
Qy      181 AGCGCAACCTATCCCAAGGTCGCCAGCGCCGAGGCGAGGCGCTACGCCGCG 240
      |||
Db      510 AGCGCAACCTATCCCAAGGTCGCCAGCGCCGAGGCGAGGCGCTACGCCGCG 569
      |||
Qy      241 TACGCTGGCCCTCTATGAGCAAGAGGATGGGGTGGGACAGATGCGCTCTGCAACC 300
      |||
Db      570 TACGCTGGCCCTCTATGAGCAAGAGGATGGGGTGGGACAGATGCGCTCTGCAACC 629
      |||
Qy      301 CGTGCTCCCGGCTGATGTTGGGCGCCCACTGACCCCGGCGTATGTCGCGTAA 360
      |||
Db      630 CGTGCTCCCGGCTGATGTTGGGCGCCCACTGACCCCGGCGTATGTCGCGTAA 689
      |||
Qy      361 AAGCATGATGATCCCTGACATGCGGCTTGCAGCACTGATGGGGTACATTCGCGT 420
      |||
Db      690 AAGCATGATGATCCCTGACATGCGGCTTGCAGCACTGATGGGGTACATTCGCGT 749
      |||
Qy      421 GCGGCTCCCTTAAAGGGGCGCTGCGAGGGCCCTGCGCAGTCCGCGGTTCTGAGAG 480
      |||

```

AHGILRACMLVRKVGAGHYLQMAFMKLAALTSTYVYDHLPLRDMWAGHGLRDLAAVE
 PVFSPDMERKILTWKADTAAAGDIIISGLPVSRBRREILIGPADSLGGGMYHLAPIT
 AYAQOTRGLGCIITSLTRDRKQVBEGRQVYSTAQSLACATCNVGCMTYVHAGSG
 ALAQGKQVTOYITNDODLVGMQAPPGKASLIPCTCGSSDLYLTRADVIVPRRG
 DSRGSLSPRPVSLKSGSGPLCPGSAVAVI FRAVCTRVAAKAVDVAPVESMETT
 MRSPIFTDMSPPAVQOTFOVAHLAPGSGSTKVPAAVAGQYVLAIVLNSVAAITL
 GFGVYMSKAGVDPNRTGRTITTTGAAPITTYGVLGADGGCGGAGAYIILICDEBS
 TDSITIGITGLDQAEAGARLVLAATTPPGSVVPHPNIEVLAISNTGELPFGK
 AIPETIKGRHLIRHSGKKKDELAALASALGNAVAAYRBLDVSIVITSGDVVYVA
 TDLMTFTGFDSDVYIDCNTVQIVDVELDTFTLETTIVQDVAVSQRKRTG
 RMGIYRVPVPRPSGMPDSVLCBYDAGCMYELTETVSRVLAAYLNTPELPCVCO
 DLHEPESVFTGLTIDAFISQTOAGCNFPYLVAQVAVCAAPPSDDOMMKC
 LRLPETHGPEPLIRYGLAONEVLTHTPIKYIMACMSADLEVSTWVYGVGLA
 ALAAYCTICVIVIGRIISGKPAVTPRODLYOEPDEMERCAASHLPIDEGMOLAB
 QPKOVALGILQITATQAEAAAVVESKRALETTPAKKRALETISGIVYAGLSTLPGN
 PAIASLMAFTASITSPLTQHTLFRNLLGAVAAQAPPSASAPVAGACIAGAVGSI
 GLGKLVLDILAGYAGVAGALVAFKMGSGEMSTBDLVNLPALISPLVGVVCAAI
 ILRRYVGPBGAVQMMNRLIAPASGKNVSPHYVPESDAAARVQIILSLITDOLK
 RLHOWINEDCSTPGSGMLRDVMDWCAVLTDFTKMLGSKLPLRPGVFFSCORGYK
 GIMRGDGIHTTCPGAQITGHVKNKSNRIVGPKTCSNMTGTPPPIAVTTGCPSP
 APNYSALMRVAAEVEYTRVGDHYVYGMTTDDNVKCPQVAPRPFTEVDGILHR
 YAPACKPLIREVITFOVGLNOYLVSQULPCEPDPVAVITSLMDTBSHITAEAKRL
 AAGSPSLASSASQLSASLAKATCTTRHSDPADLITENILMRDMEGNITRVSEN
 KVIILDFEPLQAEDEBDEISVAABILKRSRFPAMPVPWAPADPNPLBESKQEDY
 VPPVHGCPLPTKAPPIPPPRKRTVLTSTVSALALATKTFGSESSAVDSGT
 ATAPPOASGDGAGSDVSVSMSPLEGEPRDPLSDGMSVSEASESDVSCMS
 YTMGTALITPCAAESKRLPINLSNLSLPHNMVYATTRSASQOKKYTPRMQVLD
 DHYRDLKEMKAKASTVAKLISVEBACULTPSPHRSKFGYAKOVRNLSKATHT
 RSVMDLLEDSPTIDITYMANNEVFCQPEKGRGAPALIVPDLGVSEEMALYD
 VSTLPEAVMGSPSYGFOYSPGOVEFLVAVMSKXPMGFADTRFSDVTEVNDIRV
 EBSIYOCODLAEAROAISLTERLIYIGPLFNSGONCGYRGRASGLTISCGNTI
 TCYKASACRAKAKODCTMLYCGDVIYAGSAGOGNDAALRVTEMTYTSAPRG
 DDPQPEYDELITSCSNVSVYHADSGRVYIITDPTPLRAAMETARHPNYSWL
 GSIINAPLTMARMLIMTHFSIILAOBELRALDOITGACTSLELDPOITQRLH
 GNASLSHSYSGEINRVASCRLKGVPLRWRHRAVRKALISQGRAPATCGYVL
 FMVAATKTKLTPIPASHLDSGWFVAGSGDIYHSLRBRPRLMWCILLISGV
 GYLLPNR"

Db 750 GGGCCCCCTAGGGGGCGCTGCGAGGGCCCTGCGCATGCGCTCCGGGTTCTGAGAGAC 809
Qy 481 GCGTGAATATGCAACAGAGAAATTAACCGGTTGCTTTCTATCTTCCCTTGAGT 540
Db 810 GGGGTAACTATGCAACAGAGAAATTAACCGGTTGCTTTCTATCTTCCCTTGAGT 869
Qy 541 TTGCTGCTGTTTGAACAATCCAGCTTCCGTTATGAAATGAGGCAACGTCGCGGATC 600
Db 870 TTGCTGCTGCTGTTGAACAATCCAGCTTCCGTTATGAAATGAGGCAACGTCGCGGATC 929
Qy 601 TACATATGCAACAGATGCTTCAACTCAAGCATGTGTAGAGACAGCGGACATGATC 660
Db 930 TATCAGGTACCAACAGATGCTTCAACTCAAGCATGTGTAGAGACAGCGGACATGATC 989
Qy 661 ATGCAACCCCGGGGTGTGTCCTGTCTCCGGAGAGATTAATCTCCGCTGCTGGGTA 720
Db 990 ATGCAACCCCGGGGTGTGTCCTGTCTCCGGAGAGATTAATCTCCGCTGCTGGGTA 1049
Qy 721 GGGCTACTCCGACGCTGCGGGCAAGAGACGCAAGCATCCCACTGCGCAATACGACGC 780
Db 1050 GGGCTACTCCGACGCTGCGGGCAAGAGACGCAAGCATCCCACTGCGCAATACGACGC 1109
Qy 781 CACGTGATTTGCTGTTGGGGCGGCTGCTTCTGCTCCGCTATGATAGTGGGGATCTC 840
Db 1110 CACGTGATTTGCTGTTGGGGCGGCTGCTTCTGCTCCGCTATGATAGTGGGGATCTC 1169
Qy 841 TCGGATCTGTTTCTGCTGCTCTGAGTGTCACTTCCGCTGCTGCGGACATCAGAC 900
Db 1170 TCGGATCTGTTTCTGCTGCTCTGAGTGTCACTTCCGCTGCTGCGGACATCAGAC 1229
Qy 901 GTACAGAGATGCAATTTGTTCAATCTATCCCGGCAAGTACAGGTCAACCGCATGCTTG 960
Db 1230 GTACAGAGATGCAATTTGTTCAATCTATCCCGGCAAGTACAGGTCAACCGCATGCTTG 1289
Qy 961 GATATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
Db 1290 GATATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1349
Qy 1021 CCACAGAGTGTGTTGGAATGTTGGAATGTTGGAATGTTGGAATGTTGGAATGTTGGAAT 1080
Db 1350 CCACAGAGTGTGTTGGAATGTTGGAATGTTGGAATGTTGGAATGTTGGAATGTTGGAAT 1409
Qy 1081 TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
Db 1410 TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1469
Qy 1141 GTTGAAGGGGAACTTACACAGAGGGGGAACACAGGGGCGCGCCGACGCGGCTTACA 1200
Db 1470 GTTGAAGGGGAACTTACACAGAGGGGGAACACAGGGGCGCGCCGACGCGGCTTACA 1529
Qy 1201 TCCCTCTTACACCTTGGGCGGCTCAGAAATTCAGCTTGTAAACCAACGAGAGCTGG 1260
Db 1530 TCCCTCTTACACCTTGGGCGGCTCAGAAATTCAGCTTGTAAACCAACGAGAGCTGG 1589
Qy 1261 CACATCAACAGAACTGCTTGAACCTGCAATGACTCCCTTCAACCTGGGTTCTTGGCGCG 1320
Db 1590 CACATCAACAGAACTGCTTGAACCTGCAATGACTCCCTTCAACCTGGGTTCTTGGCGCG 1649
Qy 1321 CTGTTCTACACGCAAGGTTCAATGCTCCGATGCTCAAGACGCACTGGCCAGCTGCCG 1380
Db 1650 CTGTTCTACACGCAAGGTTCAATGCTCCGATGCTCAAGACGCACTGGCCAGCTGCCG 1709
Qy 1381 CCCATTTGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440
Db 1710 CCCATTTGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1769
Qy 1441 GACCAAGGCGCTTATTTGCTGGAACCTCAACCTGATGATGATGATGATGATGATGATG 1500
Db 1770 GACCAAGGCGCTTATTTGCTGGAACCTCAACCTGATGATGATGATGATGATGATGATG 1829
Qy 1501 CAGGTGTGTGGCCAGTGTACTGTTTCACTCCAGGCGCTGTTGTGTGTGGGAGCAACGAT 1560
Db 1830 CAGGTGTGTGGCCAGTGTACTGTTTCACTCCAGGCGCTGTTGTGTGTGGGAGCAACGAT 1889

Qy 1561 CGTTTCGGCGCCCTACGTAAGATGGGTGAGATGAGACGAGCTGCTTCTCAAC 1620
Db 1890 CGTTTCGGCGCTTCTACGTAAGATGGGTGAGATGAGACGAGCTGCTTCTCAAC 1949
Qy 1621 AACACGCGCGCCGCAACGAGGGAATGTTGCTGCTGATGATGATGATGATGATGATGATG 1680
Db 1950 AACACGCGCGCGCTGAGCAATGTTGCTGCTGATGATGATGATGATGATGATGATGATG 2009
Qy 1681 ACCAAGAGTGTGGGGGCGCGCGGCAACATCGGGGGGGTTCGGCAACCACTTGTATC 1740
Db 2010 ACCAAGAGTGTGGGGGCGCGCGGCAACATCGGGGGGGTTCGGCAACCACTTGTATC 2069
Qy 1741 TGCCCCAGGAGCTGCTTCCGGAAGATCCGAGGCACTTACACCAATGCGGTTGCGGG 1800
Db 2070 TGCCCCAGGAGTGTGTTCCGGAAGATCCGAGGCACTTACACCAATGCGGTTGCGGG 2129
Qy 1801 CTTTGGTTG 1809
Db 2130 CTTTGGCTG 2138

RESULT 10
AF207766
LOCUS AF207766 9379 bp RNA linear VRL 27-APR-2000
DEFINITION Hepatitis C virus strain MD25 complete genome.
ACCESSION AF207766
VERSION AF207766.1 GI:7650249
KEYWORDS
SOURCE
ORGANISM
Hepatitis C virus
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE
1 (bases 1 to 9379)
Nagayama, K., Kurosaki, M., Enomoto, N., Miyasaka, Y., Murano, F. and
Sato, C.
Characteristics of hepatitis C viral genome associated with disease
progression
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 9379)
Nagayama, K., Kurosaki, M., Enomoto, N., Miyasaka, Y., Izumi, N. and
Sato, C.
Direct Submission
Submitted (23-NOV-1999) Second Department of Internal Medicine,
Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-ku,
Tokyo 113-8519, Japan
FEATURES
source
Location/Qualifiers
1..9379
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/strain="MD25"
/db_xref="taxon:11103"
/note="Isolated from a patient with liver cirrhosis and
hepatocellular carcinoma (HCC-5)"
330..9362
/codon_start=1
/product="polyprotein"
/protein_id="AA05996.1"
/db_xref="GI:7650250"
/translation="MSTNPKFQRTKENTNRPODVKPPGGQIVGAVLLPRGRPL
GVATRTKTSERSQGRROPVPAKREGRTPAQVPWPVYEGMGAMGLSPG
SRPMPGTPDPRRSRLNGKVIDLTCGSHADPQYPIVGAALGAAALAHGVLBD
GVNATGNLPCSSIFLLALSLCTIPASAYEVNAGVHVNDCNSIIVYBAD
MIMTPGCVPCVRNNSRCMVLTPTLAAANSPVTKTIRHVDILVGAAPCSAMY
VGDIQGVFLVSQLPFTSPRRHETVODCKSIYGHLSGRHAMDMMNMSPTALIV
SOLRLIDAVMDVAGAHWGLAAYSMVGNAAKXVLIWLLPAGVDSFVSGGTA
GRTVGLARFPTSGPSONIQLVNTNGSMHIRTALNCDSIQTFIALFTHKENDS
GCPERMASRSRIDKEQGWBPITYAGSHLDORPYCMHYAPQPGIYPAAGVCPVYC
FTSPVIVGTSDRSGVPTVMGVNVSIVLLNNTRPQGMVFGCTNNNSGFTFTCGG
PCNIGVGNNTLLCPDPCRKHBEATYTIKCGSSPWLTPRLVDYVYPRIMHYPCVNF
TTLKRVNYGGEHRLNAACMWTREGCNEDRSELSPLLSITBQIILPCSTTL
PALSTGLIHQNVNDVQVLYIGISAVSVFVIKWEYVLLFLIADARVCACLMMMLL
IAQEAALLEKLVINAAISVAGHGLIFLVFCAAWYIKKLIVGAAVYALYGVPLLL

LLALPLAAMREMAASCGAVFVGLVLLTSLSPHYKFLARLIMWLITLREAL
LOYWPEPLINRGRDAILILLTCAHBEPLIDISKILLIITLPIVLAOGITRPPYPR
AQLIRKMLYRKVAGHYVQALMLKALITLTVYDHLITPLMDMAGRLDLAVE
PVLPSDEKTVITWADTACVNIISGLPVSAARGKEILLGPDHSEGRMLLAPPT
AYCOQTRGLGCIITSLTGRDKNVEGVNSTATOSFLATCMLNVCNIVYASGK
TLAGKPIITOMYTNVODLVGMWAPGASLTPCTCGSDLLTGLVTHADVITPVRG
DSRGSILSPRPVYLKSGSGGLCPGSHVGVIFRAVCTRGAKAVDPVPMEMT
TRSPPTDSSPAVPOTPOVAILHAPTSGKSTKVPAAVAGVYVIVPMVAATL
GREGVMSKAGVDPNIRITGRITTTGAPITTYTGKFLADGCGSGAYDINKDECHS
TDSITSLIGITVLDQAEVAGARLVATAPPGSVVPHNIEILSNSTGELIPFYK
ALPIEALKGRLIFCHSKKDELAASKGLANAVAYRGLADVITPVGSVVVA
TALMTFTGDLSDVIDCNTCVQTFDLSDFITLETITTPDASRSRORRTEBP
RRIYRFTVPEBPSGMPSSVCECDGACAWELTPATSVRLAVALYNTPELPYCO
DHLEFESVFTGLTHIDAHLSOTKQAGNPLYVAOATKADAPSPSMOMKIC
LIRKPTLSPPTLIRLGAONEVILTHITRTYIMACSADILEVITSVLVGGLA
ALAAVCLTGSVIVIRILLISGRPAVPDEVLVIREDEBECAHLPIEQMOALAE
OFKOKALGLOTASKOAEAPVYESKMLAEAFKHMNPIISGLOYLAGSLTEN
PATASLMAFTASITSLPTQSTLLFNILGMAVLAAPASAPFVAGIAGAVSI
GKGVLDILAGVAGVAGALVAFKVMGSEMPSTEDLVNLPALISPGALVGVCAA
ILBRHYGREGANVOMNRLIAPASRGNHSPHYVPESDAAVAVTQILSLITTLQK
RLHOMINEDSTPSCSWLMDVMDITCTVITDRKTVLQSKLRLPLVPLSCQKIK
GWRGDIHMTTCPCGAQITGHVKNMSMRVGRPTCSNTHGTFPVNAYTGTCTBP
APVYSKALMRVAEEVYEVTVGDVHYVTGMDTVNCPQVPAPEFTEVDVRLHR
YAPACKPLREEVTFQVGLNOYLVGSQLPCEPEPDVAVLITMLTDSHITAETAKRL
ARGSPSLASSASOLASBLKATCTTRHSDPADLILENLMRORMGNITPVESEN
KVITLDSFEPRLAREDBREVSVAEVLRSRKEPPALPIYAPDPNPLLESKDDY
IPVHVGCPLPPTKAPPIPPRRKRTVLTSTVSSALAEALATKRTVEASDVSNDG
ATABPQPDPAAGDVBESVSMPLBEGPDPDLSDMSWTSVEASDVSNCMS
YMTGALITPCAESKLPINALNSLIRHNNVATTSASQORKATYFDLOVLD
DHYRGLKEIKAKASTYKAKLSVEEAKLTPHARSARSKYKQAVRNISGAVNHI
BSVMKDLLEDETPIIDITTIMAKNEVFCVQPEKGRKPARLIVPDLGVYCEMALYD
VUSTLPOAWGSSYGFQYSGQVREFLVNAKSKRPMGCSYVTRPDSVITSDRT
EBSITOCDDIAPBARQIIRSLRYVGGELTNSKQNGCYTRCRASVYDTSNGTL
TCVLKSAACRAKIDCTMLVCGDVLVYCESAGQDEAASIRYETEMARTSPAG
DPQPEYDELITSCSNVSVADHAGSKVIRYTRDPTPLAAVETASTRVSNML
GNIMYAPTIMAMITMTHTFSLIAOBEKALDCOIGACISIEPLDIPOLIOHL
GSAFSLHSYSPGEINRVAISCKRKLGVPLRVRRLRARSRAULISQGRRAATCGVL
FNVAVTKLKTPIIPAASOLDLSNMFVAGISGDIYHLSRARPFRMCLLILSIVG
GYLBNR

ORIGIN

Query Match 87.4%; Score 1580.2; DB 14; Length 9379;
Best Local Similarity 92.1%; Pred. No. 0;
Matches 1666; Conservative 0; Mismatches 143; Indels 0; Gaps 0;
QY 1 ATAGACAGATCTTAAACCTCAAGAAAAACCAAGTAAACCAACCGCGCCACAG 60
Db 330 ATAGACAGATCTTAAACCTCAAGAAAAACCAAGTAAACCAACCGCGCCACAG 389
QY 61 GAGCTCAAGTCCCCGGGGGGGTGTCAGATGTTGGTGAAGTTTACTGTTCCGCCAGG 120
Db 390 GAGCTCAAGTCCCCGGGGGGGTGTCAGATGTTGGTGAAGTTTACTGTTCCGCCAGG 449
QY 121 GGGCCAGGTTGGGTGCGCGGACCTAGGAAGACTTCGAGCGGTGCAACCTCGTGA 180
Db 450 GGGCCAGGTTGGGTGCGCGGACCTAGGAAGACTTCGAGCGGTGCAACCTCGTGA 509
QY 181 AGGCGCAACTATCCCAAGGCTGCGCAAGCCGAGGCGAGGCTTGAGCTGACCCGCG 240
Db 510 AGGCGCAACTATCCCAAGGCTGCGCGGACCTGCGGCGAGGCGAGGCTGACCCGCG 569
QY 241 TACCTTGGCCCTCTATAGCAAGAGGATAGGGGTGGGCAAGATGCTCTGTCAACC 300
Db 570 TACCTTGGCCCTCTATAGCAATGAGGATAGGGGTGGGCAAGATGCTCTGTGCGCC 629
QY 301 CGTGGCTCCGCGGCTAGTGGGCGCCACTGACCCCGGCGTGGGTGCGAGTAATTTGGGT 360
Db 630 CGTGGCTCCGCGGCTAGTGGGCGCCACTGACCCCGGCGTGGGTGCGAGTAATTTGGGT 689
QY 361 AAAAGCATGATACCTTCACATGCGGCTTGCGGACCTCATGGGGATACATTCGCGTGTG 420
Db 690 AAAAGCATGATACCTTCACATGCGGCTTGCGGACCTCATGGGGATACATTCGCGTGTG 749

QY 421 GGGCTCCCTTAGGGGGCGCTGCGAGGGCCGTGGCGATGCGGCTCCGGGTTGTGAGGAC 480
Db 750 GGGCCCCCTTAGAGAGCGCTGCGAGGCCCTGGCGATGCGGCTCCGGGTTGTGAGGAC 809
QY 481 GGGGTAACTATGCAACAGGGAAATTTAACCGGTTGCTTTCTCTATCTTCTCTTGGCT 540
Db 810 GGGGTAACTATGCAACAGGGAACTTGGCGGTTGCTTTCTCTATCTTCTCTTGGCT 869
QY 541 TTGCTGCTGTTTACCATTTCCAGCTTCCGCTTATGAAAGTGGCAAGTGTCCGGATC 600
Db 870 CTGCTGCTGTTTACCATTTCCAGCTTCCGCTTATGAAAGTGGCAAGTGTCCGGATC 929
QY 601 TACCATGTCAGGAACGATTGCTCAACTCAAGCATGTGTGAGAGAGAGGACATGATC 660
Db 930 TACCATGTCAGGAACGATTGCTCAACTCAAGCATGTGTGAGAGAGAGGACATGATC 989
QY 661 ATGCAACCCCGGGGTGTGTGCTGCTGTCGCGGAGGATTAATTCCTCCGCTGCTGGGTA 720
Db 990 ATGCAACCCCGGGGTGTGTGCTGCTGTCGCGGAGAGAACTCTCCGTTGCTGGGTG 1049
QY 721 GGGCTCACTCCAGCTTCGCGGCAAGAGACGCCAGATCCCACTGCGACATACGACGC 780
Db 1050 GGGCTCACTCCAGCTTCGCGGCAAGAGACGCCAGCTCCCACTAAGACATACGACGT 1109
QY 781 CAGCTGATTTGCTGCTGGGGGCGGCTGCTGCTCCGCTATGATAGTGGGGGATCTC 840
Db 1110 CAGCTGATTTGCTGCTGGGGGCGGCTGCTGCTGCTGCTATGATAGTGGGGGATCTC 1169
QY 841 TGGGATCTGTTTCTGCTCTCTCAGCTGTTCACTTCTGCGCTGCGCGACATACGACG 900
Db 1170 TGGGATCTGTTTCTGCTCTCTCAGCTGTTCACTTCTGCGCTGCGCGGATGAGAG 1229
QY 901 GTCAGAGCTGCAATTTGCTCAATCTATCCCGGACGATACAGGTACAGCATGCTTGG 960
Db 1230 GTCAGAGCTGCAATTTGCTCAATCTATCCCGGACGATACAGGTACAGCATGCTTGG 1289
QY 961 GATATATATGATGATGCTGCTCACTCAAGAGCCTGATGATGAGGATGCTGCGGATC 1020
Db 1290 GATATATATGATGATGCTGCTCACTCAAGAGCCTGATGATGAGGATGCTGCGGATC 1349
QY 1021 CCACAGCTGCTGTCGATGATGTCGCGGGGCGCCACTGCGGAGTCTGCGGCGCTTGGC 1080
Db 1350 CCACAGCTGTCATGATGATGTCGCGGGGCGCCACTGCGGAGTCTGCGGCGCTTGGC 1409
QY 1081 TACTATTCATGAGTGGGAACTGGGCTAAGTCTTGGTGTGATGCTACTTCTTGGCGG 1140
Db 1410 TACTATTCATGAGTGGGAACTGGGCTAAGTCTTGGTGTGATGCTACTTCTTGGCGG 1469
QY 1141 GTTGAAGGGAACTTACACAGACAGGGGGGACACAGCGCGCGCCGACCGGCTTACA 1200
Db 1470 GTTGAAGGGAACTTACACAGACAGGGGGGACAGCGCGCGACCGGCTTACGCA 1529
QY 1201 TCCCTCTTACACCTGGGCGGCTCAGAAATTCAGCTTGTAAACACAAACGACGCTGG 1260
Db 1530 CGCTTCTTACACCTGGGCGGCTCAGAAATTCAGCTTGTAAACACAAACGACGCTGG 1589
QY 1261 CACATCAACAGAACTGCTTGAATGCAATGACTCTCCCTCAAACTGGGTTCTTGGCGG 1320
Db 1590 CACATCAACAGAACTGCTTGAATGCAATGACTCTCCCTCAAACTGGGTTCTTGGCGG 1649
QY 1321 CTGTTCTTACACAGGTTCAATGCTGCGGATGCTGAGACGATGCGACGCTGCGG 1380
Db 1650 CTGTTCTTACACAGGTTCAATGCTGCGGATGCTGAGACGATGCGACGCTGCGG 1709
QY 1381 CCCATTGACCAATTCATGAGGGGTGGGCTCCCATCATCTTATATGATGCCAGGCTTG 1440
Db 1710 TCTATTGACCAATTCATGAGGGGTGGGCTCCCATCATCTTATATGATGCCAGGCTTG 1769
QY 1441 GACAGAGGCTCTATTGCTGGCACTACGCACTTCAACCGTGTGTATCTGTCGCGGCTTG 1500
Db 1770 GACAGAGGCTCTATTGCTGGCACTACGCACTTCAACCGTGTGTATCTGTCGCGGCTTG 1829
QY 1501 CAGGTGTGTGGCCAGTGTACTGTTTCACTCAAGCCCTGTTGTGTGGGAGACAGCCGAT 1560

|||||
1830 CAGGTGTGTGTCAGATGATGCTTCAACCCCAAGCCCTGTTGTGTGGGACGAGTAT 1889
QY 1561 CTTTTGGGGCCCCCTACGACATGGGGTGAAGATGACACGACGTGCTTCTCAAC 1620
Db 1890 CCGTCGGGGGCTCCCTACGATTAACGGGGGGGAAAGAGTCCGACGTGCTCTCAAC 1949
QY 1621 AACACGGCGCGCCGACAGGGGCAACTGGTTCGGCTGTACATGGATGAAATGACACGGGGTTC 1680
Db 1950 AACACGGCGCGCCGCGACGGGCAACTGCTTCGGCTGTACATGGATGAAATGACACGGGGTTC 2009
QY 1681 ACCAAGACGTGTGGGGGGCCCCCGTGCACATCGGGGGGGTCCGCAACAACCTTTGATC 1740
Db 2010 ACCAAGACGTGTGGGGGGCCCCCGTGCATATCGGGGGGGTCCGCAACAACCTTGATC 2069
QY 1741 TCCCCCAGGACGTGCTTCGGGAAGACATCCCGAGGCCACTTAACCAACCAACCGGCTTCGGGG 1800
Db 2070 TCCCCCAGGACGTGCTTCGGGAAGACATCCCGAGGCCACTTAACCAACCAACCGGCTTCGGGG 2129
QY 1801 CCTGTGTTG 1809
Db 2130 CCTGTGTTG 2138

RESULT 11
AB049091 9547 bp RNA linear vRL 22-AUG-2002
LOCUS AB049091
DEFINITION Hepatitis C virus gene for polyprotein, complete cds,
ACCESSION AB049091
VERSION AB049091.1 GI:11559448
KEYWORDS
SOURCE
ORGANISM
Hepatitis C virus
Hepatitis C virus
Virus; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.

REFERENCE
AUTHORS
1
Takahashi, K., Iwata, K., Matsumoto, M., Matsumoto, H., Nakao, K.,
Hachihara, T., Ohta, Y., Kanai, K., Maruo, H., Baba, K., Hijikata, M. and
Mishiro, S.
Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients
with hepatocellular carcinoma: the 'progression score' revisited
Hepatol. Res. 20 (2), 161-171 (2001)
11348851
2 (bases 1 to 9547)
Mishiro, S.
TITLE Direct Submission
JOURNAL Submitted (19-SEP-2000) Shunji Mishiro, Toshiba General Hospital,
PUBMED Department of Medical Sciences, 6-3-22 Higashi Oh-1, Shinagawa-ku,
AUTHORS Tokyo 140-8522, Japan (E-mail: shunji.mishiro@po.toshiba.co.jp,
Tel:81-3-3764-8981, Fax:81-3-3764-8992)
JOURNAL Location/Qualifiers
FEATURES
source
1..9547
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/isolate="HCV142"
/specific_host="human"
/db_xref="taxon:11103"
/tissue_type="serum"
/note="HCV-genotype: 1b"
<1..259
/note="partial"
260..9292
/codon_start=1
/product="polyprotein"
/protein_id="BAB18604.1"
/cd_xref="GI:11559449"
/translation="MSTNPKPOKTKRNTNRPPQDVKFGGQIVGVLLPAPRL
GVATRKTSEROPRGROPIPKAROPESRAVAGYPMPLVNGSLGAGMLSPRG
SRWSGPTDPRRSRNIGKVIDTLTCGFADLMGYIPIVGAPIGCAAKLAHGVRLD
GVNATGNIPGSPSIFLALISCLTIPASATVENVSGIHTYNTDNCSSSIYIEAD
VTMAHAGVCPVRENSSRCWALPTTLAARNSVPTTTRRHVDLIVGTAAFCSAMV
VDLGGSVFLISQLFTFSPRRHETVQSCNSLYPGHIVGHRMAWMNNWSPALVY

SQILRIPOAVMDVAGAMWGLAIAAYSNVGNNAKYLIVMLIFAGVGGTHTVTCGAQ
GSAFRLTSPFTSGASOKIOLINTNGSMHINRILTNCDLSHTFLAIFTHRFNLS
CSBERLSASRSDIPDGMKPIITYAOSPSRPORYCMHVAIRPGIIVASVCGPVC
PTSPVYVGTTRDRCVPTYSNGENETDVLINNRPPQGNRCCTMNWSTGPTTCGG
PPCNIIGAGNNTLTCPTDCFRKPEEATYTKGSGPPMPLPRCMVPTPLRMHPTCTV
TIPIVRYVYGVGERHLNAAACMTGECRCNEDRSELSLPLLSSTTGMQLPSCFNP
PALSTGILHLQNVNVDQVYLIVGSAFVSVLFEKVEYLLFLFLADARVACIMMLL
IAOAEALENLVYNAAASVAGAGIISLVFFCSAMVYIKGLVGAAYAFEGWPLR
LILALPRAYAMDREMASSCGGCVFVGVLLTSLPYKVLPLAKLIMLOYITPABH
LQWVPLPRNRGSDAILILTCANPBLIPITITLLALIRGLMVLQAGITRVYFPR
AQLIRACMLVRKAGGHTVOMAFKLAALDGTIVYHLPFLRMATGLDLVAAYE
PVFSDMEYTKLIWGAADTAAGDILISGLPVSARGRRLILGAPDSLEGQKRLAPLT
AYSQOTRGLGCIITSLTRDXNVEBSVQVASTATOSFLATCVNGCMVYHRAGSK
TLAPKPEITPMYTNVODLVGMQAPGARSILCTCGSSDLVLTTHADVIYPRRG
DSRGLSPRPVYIKGSSGGPILCPGHWVIGIRAAVCTRGVAKAVADVPEAMET
MRSVPTDNSPVPVPTOTFOVAHLAARPSKSTKQVAAVAAQCVKYLVINPSAATL
FGAYMSKAGHTDNPINRTGIRITTTTGAPIYTSYTKVLADGGCGGAGDITICEBS
TDSPTTILGIGTVALDQAEFAGARLVLTATPSPSVVPHNPIEVALSNGETIPEYK
AIPIEAIKGGHILFCHSKKKCDELAKLSSGINAAVAYNGLDVSVIPISGNVVVA
TDALMTGYTGDPDIDNCTCVTQVDFSLDPTPTIETTVVPODASRTORGRTRG
RGIYRFVTPGERBSGMDSSVLCEYDAGCAMYELPAETSVRLRAYLNTGSLPVC
DHLFEMESVPTGLTHIDAHPLSOTKQAGNDPVLYAOATVYCARAQAAPPSMDMFC
LIRKPLTHGPTPLYLKGAONSVTLTHPTTKIYACMADELEVSTVYVGVYLA
ALAAVCLTGSVIVIGRIILSGRAVILPDEVLTRFDEMEBCASHLYIEQMQLAB
QFKOKALGLQOTATKQAEAAAPVESKMRALAEFAWMNMFISGIYLAGLSTLPGN
PAISALMAFTASVTSPLTQSTLFLNIGVAAQLAPPSAASFVAGIAGAAGVSI
GLKVLVIDIAGVAGVAGALVARKMSGEMSTEDLVNLLPALISGALVGVCAAI
ILRHVNGRGAOVOMNRILAFASRGHWSPHYRVPSDGAARVTOILSLITPOLK
RLHWINEDCSTPCSGMWLRDWMTCVILIDFTWLOSLKPLRBPVPFSCORRYK
GVMRGDIDIMHTTCCGAQIAGHVANKSIRLVGPTKCTSNWHTGPIINAYTGPCTPSP
APNYSRALMRYAEVYEVTVGVYHYITGATTDNIPCQVPAPEPFEVDGVRLLR
YAPACKPLIREAEVQVNLQVYLSOLPCGEPEDVAVLITSMITDPSHILYETAKRL
ARGSPMASSASOLSAPSLKAACTTHSSPDLLLEANILMEKGNATRVESN
KVYILDSFDPILAREDERVSPARILIKKSNPPAMPINARPDYNPILSEMDPY
VPTVHGCPLDPTTAPPIPPRRKRTVYVLESTYSALAEATTFESSBSASDGT
ATAPDASIDDDKESVESYSNMPLEBEGDDLSLDSSTVSEAEATVSCDSMS
YTWGALITPCAAESKLPINALNSKILRHNNVYATTSRSASLRQKVFDRQVLD
DHYVADIKEMEKAKASYEACKPLTPPSASKEFYGADVNLSSKAINHI
RSVKDILDEDTEPIIDITIMAKNVPFCQPKGRKARLIVFDLGRVCKEKNALY
VNSTLPOAVMSSYGRQYSPGRVREPLVYKTMKSKCKMGSGESYDTRCTENDIV
EBSITVOCDDIAPBRQAIRSLTEFLYIGGPTMNSKQNCQRRRBRASGVLTSCGNTL
TCYIKATPAACCAALQDCTMLVNVDLIVICESGTOEDANLLEFTEAMRYSPAGP
DPPQEVYDELITSCSSNVSAHADSKRKYVLLTRDPTLSRAEMTARTPNSWL
GIMYIAPTLIARMILMTHFPSILAOOLEKALDOICYGASIEPDLIOLIERHL
GLSAFSLHSYSGEINRVSCLRLGVPPLVWHRARSYAKLLSOGGRATCGKYL
FMAVVRKTKLTIPIASQDLDSGFVAGISGDIYHLSARBRMFLCILLISGV
GTYLLENR"
9293..9331
/note="upstream of poly-U stretch"
9332..9425
/note="poly(U) length is indistinct
poly(U) stretch"
9426..9547
/note="3' X tail"

ORIGIN
3'UTR
9293..9331
/note="upstream of poly-U stretch"
3'UTR
9332..9425
/note="poly(U) length is indistinct
poly(U) stretch"
9426..9547
/note="3' X tail"

Query Match 87.1%; Score 1575.4; DB 14; Length 9547;
Best Local Similarity 91.9%; Pred. No. 0;
Matches 1663; Conservative 0; Mismatches 146; Indels 0; Gaps 0;
QY 1 ATGAGCAGCAATCTCTAAACCTCAAGAAAACCAACGTAACACCAACCGCGCCACAG 60
Db 260 ATGAGCAGCAATCTCTAAACCTCAAGAAAACCAACGTAACACCAACCGCGCCACAG 319
QY 61 GACGTCAAGTCCCGGGCGGTGTGTCAGATCGTTGTGAGATTACCTTTGCCGCGCAGG 120
Db 320 GACGTCAAGTCCCGGGCGGTGTGTCAGATCGTTGTGAGATTACCTTTGCCGCGCAGG 379
QY 121 GACCCAGGTGGGTGTCGCGCGCATGAGAAAGACTTCGAGCGGTGCGCAACTCTGTGGA 180
Db 380 GACCCAGGTGGGTGTCGCGCGCATGAGAAAGACTTCGAGCGGTGCGCAACTCTGTGGA 439
QY 181 AGGCGACAACCTATCCCAAGGCTCGCCAGCCGAGGGCAGGAGCCTGGGCTCAAGCCCGG 240

```

Db 440 AGGGGAAACCTATCCCAAGCTCGCAGCCGAGGGGCAAGGGCTGGGCTCAGCCGGG 439
Qy 241 TACCTTGGCCCTCTATAGGCAAGAGGAGTGGGCGAGAGTGGCTCTGTCC 300
Db 500 TATCTTGGCCCTCTATAGGCAAGAGGAGTGGGCGAGAGTGGCTCTGTCC 559
Qy 301 CGGGGCTCGGCTGAGTGGGCGCCGAGAGTGGGCGAGAGTGGGCGAGAGTGGG 360
Db 560 CGGGGCTCGGCTGAGTGGGCGCCGAGAGTGGGCGAGAGTGGGCGAGAGTGGG 619
Qy 361 AAGTATGATACCTCTATAGGCAAGAGGAGTGGGCGAGAGTGGGCGAGAGTGGG 420
Db 620 AAGTATGATACCTCTATAGGCAAGAGGAGTGGGCGAGAGTGGGCGAGAGTGGG 679
Qy 421 GGGGCTCGGCTGAGTGGGCGCCGAGAGTGGGCGAGAGTGGGCGAGAGTGGG 480
Db 680 GGGGCTCGGCTGAGTGGGCGCCGAGAGTGGGCGAGAGTGGGCGAGAGTGGG 739
Qy 481 GGGGCTCGGCTGAGTGGGCGCCGAGAGTGGGCGAGAGTGGGCGAGAGTGGG 540
Db 740 GGGGCTCGGCTGAGTGGGCGCCGAGAGTGGGCGAGAGTGGGCGAGAGTGGG 799
Qy 541 TTTGCTGCTGTTGACATTCAGCTTCCGCTTATGAGTGGGCGAGAGTGGGCGAG 600
Db 800 TTTGCTGCTGTTGACATTCAGCTTCCGCTTATGAGTGGGCGAGAGTGGGCGAG 859
Qy 601 TACATGTCAGAGAGTGGGCGCCGAGAGTGGGCGAGAGTGGGCGAGAGTGGG 660
Db 860 TACATGTCAGAGAGTGGGCGCCGAGAGTGGGCGAGAGTGGGCGAGAGTGGG 919
Qy 661 ATGCAACAGGCGGCTGAGTGGGCGCCGAGAGTGGGCGAGAGTGGGCGAGAGT 720
Db 920 ATGCAACAGGCGGCTGAGTGGGCGCCGAGAGTGGGCGAGAGTGGGCGAGAGT 979
Qy 721 GGGCTCACTCCAGAGTGGGCGCCGAGAGTGGGCGAGAGTGGGCGAGAGTGGG 780
Db 980 GGGCTCACTCCAGAGTGGGCGCCGAGAGTGGGCGAGAGTGGGCGAGAGTGGG 1039
Qy 781 CAGCTGATTTGCTGTTGAGGCGGCTGCTTGTCTGCTGCTATGTAAGTGGGAGT 840
Db 1040 CAGCTGATTTGCTGTTGAGGCGGCTGCTTGTCTGCTGCTATGTAAGTGGGAGT 1099
Qy 841 TGGGATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 1100 TGGGATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1159
Qy 901 GTACAGAGTGAATGTTGTAATCTATCCGAGCGAGTATCAGGTACCGAGTGG 960
Db 1160 GTACAGAGTGAATGTTGTAATCTATCCGAGCGAGTATCAGGTACCGAGTGG 1219
Qy 961 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
Db 1220 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1279
Qy 1021 CCAAGAGTGTGCTGAGCATGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
Db 1280 CCAAGAGTGTGCTGAGCATGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGG 1339
Qy 1081 TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Db 1340 TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1399
Qy 1141 GTTACAGGAGGAGCTTACAGCAGAGGAGGAGCAGCAGGCGCGCGCGCGCTTAC 1200
Db 1400 GTTACAGGAGGAGCTTACAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1459
Qy 1201 TCCCTCTTACAGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1260
Db 1460 TCCCTCTTACAGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1519
Qy 1261 CACATCAACAGAGTGGCTGAGAGTGGAGTGGCTGAGAGTGGAGTGGCTGAGAG 1320

```

```

Db 1520 CACATCAACAGAGTGGCTGAGAGTGGAGTGGCTGAGAGTGGCTGAGAGTGG 1579
Qy 1321 CTGTTTACACAGCAGAGTTCATGCTGCGGATGCTCAGAGGCGCATGCTGCTGCG 1380
Db 1580 CTGTTTACACAGCAGAGTTCATGCTGCGGATGCTCAGAGGCGCATGCTGCTGCG 1639
Qy 1381 CCAATTGACAGTTCATGATGAGGAGTGGGCGGCGGCGGCGGCGGCGGCGGCGG 1440
Db 1640 TCCATTGACAGTTCATGATGAGGAGTGGGCGGCGGCGGCGGCGGCGGCGGCGG 1699
Qy 1441 GACCAAGGCGCTTATGCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1500
Db 1700 GACCAAGGCGCTTATGCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1759
Qy 1501 CAGGTGTGAGGCGGAGTGTCTGTTTCACTCCAGCCCTGTTGTGTTGAGGAGCAG 1560
Db 1760 CAGGTGTGAGTGTGAGTGTGTTTCACTCCAGCCCTGTTGTGTTGAGGAGCAG 1819
Qy 1561 CGTTTGGGCGCCCTGATGATGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGT 1620
Db 1820 CGTTTGGGCGCCCTGATGATGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGT 1879
Qy 1621 AACAGCGGCGGCGGAGCAGGAGCAGTGGTTCGCTGTATGATGATGATGATGATG 1680
Db 1880 AACAGCGGCGGCGGAGCAGGAGCAGTGGTTCGCTGTATGATGATGATGATGATG 1939
Qy 1681 ACCAAGCTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1740
Db 1940 ACCAAGCTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1999
Qy 1741 TGGCCAGAGGAGTGTGCGGAGAGTCCCGAGGCGCAGTTCACCAATGCGGTTG 1800
Db 2000 TGGCCAGAGGAGTGTGCGGAGAGTCCCGAGGCGCAGTTCACCAATGCGGTTG 2059
Qy 1801 CCTTGGTTG 1809
Db 2060 CCTTGGTTG 2068

RESULT 12
AX739971
LOCUS AX739971 9605 bp DNA linear PAT 08-MAY-2003
DEFINITION Sequence 2 from Patent WO02059321.
ACCESSION AX739971
VERSION AX739971.1 GI:30519245
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
Virus; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.

REFERENCE
1 de Francesco, R., Migliaccio, G. and Paonessa, G.
Hepatitis C virus replicons and replicon enhanced cells
Patent: WO 02059321-A 2 01-AUG-2002;
ISTITUTO DI RICERCHE DI BIOLOGIA MOLECOLARE P. ANGILETTI S.P.A.
(IT)

FEATURES
source Location/Qualifiers
1..9605
/organism="Hepatitis C virus"
/mol_type="unassigned DNA"
/db_xref="taxon:1103"
/notes="Con 1 HCV isolate amino acid"

ORIGIN
Query Match 86.9%; Score 1572.2; DB 6; Length 9605;
Best Local Similarity 91.8%; Pred. No. 0;
Matches 1661; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

Qy 1 ATGAGCAGGATCTTAACCTCAAGAAACCAACGTAACCAACGCGCGCCACAG 60
Db 342 ATGAGCAGGATCTTAACCTCAAGAAACCAACGTAACCAACGCGCGCCACAG 401
Qy 61 GACGTCAAGTTCGCGGCGGTGTCAGATCGTTGTGAGGTTTACTGTTGCCGCGAG 120

```

|||||
Db 402 GAGCTCAAGTCCCGGCGGTGTGAGATGTCGGTGAGATTACCTGTTGCCGCGCAGG 461
121 GGGCCCAAGTTGGTGTGTGGCGCGCATAGAGAAGCTTCGAGCGGTCCGAACCTGTGTGA 180
Db 462 GGGCCCAAGTTGGTGTGTGGCGCGCATAGAGAAGCTTCGAGCGGTCCGAACCTGTGTGA 521
181 AGGCGCAACACTATCCCAAGGCTCGCCAGCCGAGAGGCGAGGGCCTGAGCCTCGAGCCGG 240
Db 522 AGGCGCAACACTATCCCAAGGCTCGCCAGCCGAGAGGCGAGGGCCTGAGCCTCGAGCCGG 581
241 TACCTTGGCCCTCTATGTGGCAACGAGGCGATGGGGTGGCGAGATGCTCTGTACCC 300
Db 582 TACCCCTGGCCCTCTATGTGGCAACGAGGCGATGGGGTGGCGAGATGCTCTGTACCC 641
301 CGTGGCTCCGGGCTAGTTGGGGCCCACTGACCCCGCGGTAGAGTCGGTAAATTGGGT 360
Db 642 CGTGGCTCCGGGCTAGTTGGGGCCCACTGAGACCCCGCGGTAGAGTCGGTAAATTGGGT 701
361 AAGTCATCATACCTCATATGCGGCTTCGCGGACCTCATGAGGATACATTCGCGTCTGC 420
Db 702 AAGTCATCATACCTCATATGCGGCTTCGCGGACCTCATATGAGGATACATTCGCGTCTGC 761
421 GCGCTCCCTTAAAGGGGGCGCTGCCAGAGGCTTGGCGGATGGCGTTCGGAGAGC 480
Db 762 GCGCGCCCTTAAAGGGGGCGCTGCCAGAGGCTTGGCGGATGGCGTTCGGAGAGC 821
481 GGGGTGAATATGCAACAGGGGAATTTACCGGGTTGCTTCTATCTCTCTCTGCT 540
Db 822 GGGGTGAATATGCAACAGGGGAATTTACCGGGTTGCTTCTATCTCTCTCTGCT 881
541 TTGCTGTCTGTGTGACCAATTCAGCTTCGCTTATAGAGTGCAGCACTGTCCGGAGTC 600
Db 882 TTGCTGTCTGTGTGACCAATTCAGCTTCGCTTATAGAGTGCAGCACTGTCCGGAGTC 941
601 TACCATGTCAAGAGATGCTTCACTCAAGCATGTGTATGAGACAGCGGACATGATC 660
Db 942 TACCATGTCAAGAGATGCTTCACTCAAGCATGTGTATGAGACAGCGGACATGATC 1001
661 ATGCAACCCCGGGGGTGTGCTGTGTCGGGAGGGTAAATTCCTCCCGCTGCTGGGTA 720
Db 1002 ATGCAACCCCGGGGGTGTGCTGTGTCGGGAGAGAACATCTCTCCCGCTGCTGGGTA 1061
721 GCGCTCACTCCACAGCTCGCGGCGCAAGAGCGCAGCATCCCATCTGCACAAATACAGC 780
Db 1062 GCGCTCACTCCACAGCTCGCGGCGCAAGAGCGCAGCATCCCATCTGCACAAATACAGC 1121
781 CACGTGATTTGCTGTTGGGGGGCTGCTTCTGCTCCGTATGTACGTGGGGATCTC 840
Db 1122 CACGTGATTTGCTGTTGGGGGGCTGCTTCTGCTCCGTATGTACGTGGGGATCTC 1181
841 TGGGATCTGTTTCTGCTCTGCTCAAGCTTCACTTCCTGCTCCGCTGCGGACATCAGC 900
Db 1182 TGGGATCTGTTTCTGCTCTGCTCAAGCTTCACTTCCTGCTCCGCTGCGGACATCAGC 1241
901 GTACAGAGCTGAATTTTCAATCTATCCCGGCGAGTATCAGGTCAACCGATGGCTTGG 960
Db 1242 GTACAGAGCTGAATTTTCAATCTATCCCGGCGAGTATCAGGTCAACCGATGGCTTGG 1301
961 GATATGATGATGAATGTCGTCACCTTACAGAGCCCTATGATGTACGAGCTACTCCGATC 1020
Db 1302 GATATGATGATGAATGTCGTCACCTTACAGAGCCCTATGATGTACGAGCTACTCCGATC 1361
1021 CCACAGAGCTGTGTGACATGTGTGGGGGGGCCCACTGGGAGTCTTGCGCGGCTTGGC 1080
Db 1362 CCACAGAGCTGTGTGACATGTGTGGGGGGGCCCACTGGGAGTCTTGCGCGGCTTGGC 1421
1081 TACTATTTCATGTGTGGGAACTGGGCTAAGTCTTGTTGTATGTACTCTTTGCCGCG 1140
Db 1422 TACTATTTCATGTGTGGGAACTGGGCTAAGTCTTGTTGTATGTACTCTTTGCCGCG 1481
1141 GTTGAAGGGAACCTTACAGAGAGGGGAGACACAGGCGCGCGCCACAGGGGCTTACA 1200
|||||

Db 1482 GTTGAAGGGAACCTTATGTACAGGGGGAGCATGTGCGCAAAAAACCCCTGGGATTAACG 1541
1201 TCCCTCTTACACACCTTGGGCGGGCTCAGAAAATCCAGCTTGTAAACCAACGAGGAGCTGG 1260
Db 1542 TCCCTCTTATACCCGGGCTCATCCAGAAAATCCAGCTTGTAAACCAACGAGGAGCTGG 1601
1261 CACATCAACAGAACTGCTTGAATGCAATGCAATGCTCCCTCAAACTGGGTTCTTGGCGCG 1320
Db 1602 CACATCAACAGAACTGCTTGAATGCAATGCAATGCTCCCTCAAACTGGGTTCTTGGCGCG 1661
1321 CTGTTCTTACAGCAGCTTCAATGCTTCCGATGCTTCAAGGCGCATATGCGGACGCTGGC 1380
Db 1662 CTGTTCTTACAGCAGCTTCAATGCTTCCGATGCTTCAAGGCGCATATGCGGACGCTGGC 1721
1381 CCCATGACAGCTTGCATCAGAGGGGGGTCCCATCATCTTATTAATGAGTCCACAGGCTTG 1440
Db 1722 CCCATGACAGCTTGCATCAGAGGGGGGTCCCATCATCTTATTAATGAGTCCACAGGCTTG 1781
1441 GACCAAGAGCCCTTATGCTGGCACTACGACCTGACCCGCTGATGTCGCGCGCTTG 1500
Db 1782 GACCAAGAGCCCTTATGCTGGCACTACGACCCGCGCTGATGTCGCGCGCG 1841
1501 CAGGTGTGTGCGCCAGTGTACTGTTTCACTCAAGCCCTGTTGTGTGTGGGACGACGAT 1560
Db 1842 CAGGTGTGTGCGCCAGTGTACTGTTTCACTCAAGCCCTGTTGTGTGTGGGACGACGAT 1901
1561 GCTTTCGCGCGCCCTTACAGATGAGAGGGGTGAGATGAGACGAGCGTGTCTCTCAAC 1620
Db 1902 GCTTTCGCGCGCCCTTACAGATGAGAGGGGTGAGATGAGACGAGCGTGTCTCTCAAC 1961
1621 AACACGCGGCGCCACAGGGGCAACTGTTGCGCTGTATCATGATGATGACACGCGGTTTC 1680
Db 1962 AACACGCGGCGCCACAGGGGCAACTGTTGCGCTGTATCATGATGATGACACGCGGTTTC 2021
1681 ACCAAGAGCTGTGGGGGGCCCCCGTGCACAACTCGGGGGGGTGGGCAACAACATTTGATC 1740
Db 2022 ACCAAGAGCTGTGGGGGGCCCCCGTGTAACTCGGGGGGATCGGCAATAAAACCTTGAC 2081
1741 TGCCCCACAGAGCTGTTCCGAGAGATCCGAGGCACTTACACCAATGTGGGTTGGGG 1800
Db 2082 TGCCCCACAGAGCTGTTCCGAGAGATCCGAGGCACTTACACCAATGTGTGGGG 2141
1801 CTTTGGTTG 1809
Db 2142 CTTTGGTTG 2150
RESULT 13
HCU238799
LOCUS Hepatitis C virus type 1b complete genome, isolate Conl.
DEFINITION
ACCESSION AJ238799
VERSION AJ238799.1 GI:5420376
KEYWORDS complete genome; core protein; glycoprotein E1; glycoprotein E2; NS2 proteinase; NS3 proteinase/helicase; NS3/4A proteinase cofactor; NS4b protein; NS5A phosphoprotein; NS5B RNA dependant RNA polymerase; p7 peptide; polypeptide.
SOURCE Hepatitis C virus type 1b
ORGANISM Hepatitis C virus type 1b
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
REFERENCE 1
AUTHORS Lohmann V., Korneer, F., Koch, J., Heitman, U., Theilmann, L. and Bartenschlager, R.
TITLE Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line
JOURNAL Science 285 (5424), 110-113 (1999)
MEDLINE 99322193
PUBMED 10390360
REFERENCE 2 (bases 1 to 9605)
AUTHORS Bartenschlager, R.
TITLE Direct Submission
JOURNAL Submitted (10-MAY-1999) Bartenschlager R., Institute for Virology,

Johannes Gutenberg - University Mainz, Obere Zahlbacher Strasse 67,
55131 Mainz, GERMANY
Location/Qualifiers
1. .9605
/organism="Hepatitis C virus type 1b"
/vifion
/mol_type="genomic RNA"
/strain="RB"
/isolate="RB"
/db_xref="taxon:31647"
/country="Germany"
1. .341
/evidence=experimental
342. .9374
/codon_start=1
/product="polyprotein"
/protein_id="CAB46677.1"
/db_xref="GI:5420377"
/db_xref="GOA:Q9MWX2"
/db_xref="UniProt/TREMBL:Q9MWX2"

5'UTR
CDS

```
/translation="MSTNPKPKTKTKNTKRRKQDYKFGGCGIYVGGVYLIPRRRL
GVAIRKTSERSQPRGRKQPIPAKQPEGAMAKQEPYMPLYNEGILGAMGWLISRG
SRSPGPTDERRSRNIGKVIDTLTGADLMGYIPLVGAFLGALAHGVRVED
GVNATGNLPGCSFSIFLLALSLCTIPASAYEVNRYGVYHTDNCNASIYEAD
MIMHTPGCVCEENSSRCVVALPTLAARNASVPTTIRRHVDLLVGAALCSAMY
VGLDGSVFLVQLFTSPRRHETDQNCISYRGHTGRAMAMMMMSPTAALV
SGLRTPQAVYDVNAGHNGVLAGLAYSVNGMAKYLIMLLPAGVDSITYTGIM
AKNTLITLSFSGSQKIQLVNTGSMHINRTALCNLSLNTGFLAALFYHKESS
GCEERNASCSPIDAFAGQWGPITYNESHSDORPYCHVAHPCGIVPAQVCGPYC
FTSPVAVGTTRDFGVITYSGENETDVLINTRPPQGNMFCGTWNSGTGFKTCG
PCNIGIGIKTKLCPDCEPKRBEATYVTCGSGPMPLPRCLVHYPRYLWYCTNF
TTEKVMYGVGVLEHLEACNMTGRGRCNLEDBRELSPLISTEMOYRCSFTTL
PALSTGLIHQNVVDVYIYIGSAVNSPALKWERYLLFTLLADARCACLMMMLL
LQAEALLENVLNMAASVAGAGHLSFLVFCAMYIKGRIVGAVALYGVPMILL
LIALPPRAYVNDREMAASCGAVFGLILTLSPHYKLFARLIMWLQYFIRAH
LQWIPLVNRRGRDAVILTCALHBEILFTIKILAILGPMVLAQITKPYVR
AHLIRACMLVRKAGVYOMALMLALITGVYVDHLPLRMAHAGLDAVVE
PVESDMETKVLITGSLDRTAACDIIILGLPVSARGREIHLGPADSLGGCMRLAIT
AYSQTRGLGCIITSLTGRDRQVEGEQVYSTATOSFLATCNVGCVTYHGRSK
TLAEPKPIITOMITNVDDLVMOAPPGASLTPTCGSSDLVTRHADVLEVRGG
DSRGLSPRPVSYGLSSGGPLCPSGHAGVIFRAVCTRGAKADVPEVSMETT
MRSPTDNDSPYAVPOTFOVAHLHAPTSCKSTKVAAYAAQYKVLVNSVATL
GFCAYMSKAHGDIPNIRGTGRTITGAPITSYTGKELADGGSGGAYDIIICDECHS
TOSTTIGTIVDOAETAGARLVVATATPSPSVVPHNIEVSLSTGERPYRK
AIPETIKSGRHLIFCHSKKDELAKEISGLCINAVAYRGLDVSVITSGDIVA
TALMTGFTGDESDVIDCNVCTVQVDFSLDPTFTLETTVPQDAVRSQRRRTG
RMGIYRFTVPERSGMPSVLCCEYDGCAYELPASTVRLAYINTPGLPYCQ
DLHEFESVFTGLTHIDAFHLSQTKOAGNFPVLVAYQATVCARQAQAPPSMDQMK
LRLKPTLGPPLIYRLGAVONEVTTHTPITKYIMACMSADLEVVTSTMVVGVTLA
ALAAVCLTGSVIVIGRIITLSGKPAIIPREVIYREFDEMEBCASHLPIEOMQAL
QROKATIGLQTHKOEAAPVRESKMTLEFMAKHMNFIISGIOYLAGSLTEGN
PAIASLMAFTASTSPLOTHTLLEFNLIGWAAQAAPPASASAFAGAAVAGSI
GKGVLYDILAGYAGVAGALVAFKMSGEMSTEDLVNLLPILISGALVVGVA
ILRBHVGPEGAVOMNRLIAFASRGHNSPTHYVESPAARVQIISLTITOLK
RLOHMINEDCTPCSGSMRLDWDMICTVLTDEKTMLOKTLPRLDGVEFGQRRYK
GYMRGDMQTTCTPCGAQITGHYKNSMAIVGPTCSNTWHPFEPINATYTGCTSP
ANYSBALMRVAEEYEVTRGVDFHYVGTDDNKCCQVPAPEFEVEGVRLHR
YAPACKPILREEVTLVGLNOYLVGSLQCEPDPDAVLNMLTDSHITATKARL
AGSPSLASSSSASQLASPLKATCTTRHSDPADLEIANLIMROGMGNITRVESN
KVLILDSFEPLQAEDEBEREVSVAELIERSKPRAMPINARDVNPPLLESKDDY
VPPVHVGCP.LPAKAPPIPPRRKRTVLSSETVSALALATKTKPSESSEANSGT
ATASPDQSDDDAGSDVESYSMPLBEPDPDPLSDSKSTVSEBAEDVYCCMS
YTMGALITPCAERTKLPINALSNSLHLHNDLVATRSLSLRKKTVTRLOVLD
DHYRDLKEMKAKASTVAKLSEBAKCLTPHSAKRSFGAGKVRNLSKRAVLI
RSVMDLDEDTETPIDTITMAKNEVFCVPEKGRKPAALIVFPDGVCEKMAIYD
VSTLPOAVMGSSSYGFOYSPGOVRVEFLVNAKAKKCPAFADVTFDSTVENDIRV
ESIIYAOCCALAPBARQALSLTERLYIGPLTNSKGNCGNRYCARCSGLTNSCANTL
TCYLAKAASCRAKLQDCTMLVCGDILVYICISAGQDEBALSIRATEAMTKYSAPG
DPRKPYDLELITSCSNVSVAHDSGRKVIYUTRDPITPLAARAEIARHPVNSW
GNIINAPTLMARMLIMTHFISILLOEQLERLADCOIAGVACTISLEPDLPIORLH
GASPLSHSYPBEIINRVASCIKRLGVPLRWRHARVARRALLSGERRAATCGTYL
FMAVATTKLKLTIPIPAASQDLSMFWAGSYGGDIYHSLRARPRFMCLLLISGV
GYLLPNR"
```

	mat_peptide	342. .914	/gene="core"
	mat_peptide	/product="core protein"	
	mat_peptide	915. .1490	/gene="E1"
	mat_peptide	/product="glycoprotein E1"	
	mat_peptide	1491. .2579	/gene="E2"
	mat_peptide	/product="glycoprotein E2"	
	mat_peptide	2580. .2768	/gene="p7"
	mat_peptide	/product="p7 protein"	
	mat_peptide	2769. .3419	/gene="NS2"
	mat_peptide	/product="NS2 proteinase"	
	mat_peptide	3420. .5312	/gene="NS3"
	mat_peptide	/product="NS3 proteinase/helicase"	
	mat_peptide	5313. .5474	/gene="NS4A"
	mat_peptide	/product="NS3/4A proteinase cofactor"	
	mat_peptide	5475. .6257	/gene="NS4B"
	mat_peptide	/product="NS4B protein"	
	mat_peptide	6258. .7598	/gene="NS5A"
	mat_peptide	/product="NS5A phosphoprotein"	
	mat_peptide	7599. .9371	/gene="NS5B"
	mat_peptide	/product="NS5B RNA dependant RNA polymerase"	
	gene	342. .914	/gene="core"
	gene	915. .1490	/gene="E1"
	gene	1491. .2579	/gene="E2"
	gene	2580. .2768	/gene="p7"
	gene	2769. .3419	/gene="NS2"
	gene	3420. .5312	/gene="NS3"
	gene	5313. .5474	/gene="NS4A"
	gene	5475. .6257	/gene="NS4B"
	gene	6258. .7598	/gene="NS5A"
	gene	7599. .9371	/gene="NS5B"
	3'UTR	9374. .9605	/evidence=experimental
ORIGIN			
Query Match	86.9%;	Score 1572.2;	DB 14; Length 9605;
Best Local Similarity	91.8%;	Pred. No. 0;	
Matches 1661;	Conservative	0;	Mismatches 148; Indels 0; Gaps 0;
QY	1	ATGAGCAGCAATCTCTAAACCTCAAGAAACCAACGTAACCAACCGCGCCACAG	60
DB	342	ATGAGCAGCAATCTCTAAACCTCAAGAAACCAACGTAACCAACCGCGCCACAG	401
QY	61	GACGTCAAGTCCCGGCGGTGTCAGATCGTGTGAGATTAACTGTGGCGGCGAG	120
DB	402	GACGTCAAGTCCCGGCGGTGTCAGATCGTGTGAGATTAACTGTGGCGGCGAG	461
QY	121	GGCCCCAGTTGGGTGCGCGCGCACTAGAGAGACTTCCGACGGTCCGACCTGTGGA	180
DB	462	GGCCCCAGTTGGGTGCGCGCGCACTAGAGAGACTTCCGACGGTCCGACCTGTGGA	521
QY	181	AGCGCAACAACCTATCCCAAGGCTGCGCAGCCGAGGCGAGGCTGTGAGCTCAGCCGG	240
DB	522	AGCGCAACAACCTATCCCAAGGCTGCGCAGCCGAGGCGAGGCTGTGAGCTCAGCCGG	581

241 TACCTTGGGCCCCCTATATGGAACAGAGGCGATGGGGTGGGAGAGATGGCTCTGTCACCC 300
582 TACCCCTGGGCCCCCTATATGGAACAGAGGCGCTTGGGGTGGGAGAGATGGCTCTGTCACCC 641
301 CGTGGCTCCGGGCGCTATGTTGGGGCCCCACTGACCCCCCGGCGTAAAGTCCGCTAATTTGGGT 360
642 CTTGGCTCTGGGCGCTATGTTGGGGCCCCAGGACCCCCCGGCGTAAAGTCCGCTAATTTGGGT 701
361 AAGATCATGATACCTCATGCGGCGCTTGGCGAGCTCATGAGGGGTAACTTCGCGCTGTC 420
702 AAGATCATGATACCTCATGCGGCGCTTGGCGAGCTCATGAGGGGTAACTTCGCGCTGTC 761
421 GGGCGTCCCTTATAGGGGGCGCTGCGAGGGCGCTGGGCGAGTGGCGGCTTGGAGAGAC 480
762 GGGCGCCCCCTATAGGGGGCGCTGCGAGGGCGCTGGGCGAGTGGCGGCTTGGAGAGAC 821
481 GGGCGTAACTATGCAACAGGGAATTTAACCGGCTTGGCTTCTATATCTTCTTGGCT 540
822 GGGCGTAACTATGCAACAGGGAATTTAACCGGCTTGGCTTCTATATCTTCTTGGCT 881
541 TTGCGTCTGTTTATGCAATTCAGCTTCCGCTTATGAGATGGGGAAGTCCGAGATC 600
882 TTGCGTCTGTTTATGCAATTCAGCTTCCGCTTATGAGATGGGGAAGTCCGAGATC 941
601 TACCATGTCACGAAAGATGCTCAACTCAAGCATCTGTATGAGAGACAGCGGACATGATC 660
942 TACCATGTCACGAAAGATGCTCAACTCAAGCATCTGTATGAGAGAGCGGACATGATC 1001
661 ATGCAACACCCCGGGGTGTGCTGTGTCTGCGAGAGGTAATTCCTCCCGCTGTGGGTA 720
1002 ATGCAACACCCCGGGGTGTGCTGTGTCTGCGAGAGACACTCTCCCGCTGTGGGTA 1061
721 GGGCTCACTCCAGCGTGGGGGCGAAGAGCGCAGATCCCACTGGCGCAATAGAGAGC 780
1062 GGGCTCACTCCAGCGTGGGGGCGAAGAGCGCTAGAGTCCCACTAGAGAGATAGAGAGC 1121
781 CAGCTGATTTCTCGTGGGGCGGCTGCTTCTGCTCCGCTATGATCACTGGGGATCTC 840
1122 CAGCTGATTTCTCGTGGGGCGGCTGCTTCTGCTCCGCTATGATCACTGGGGATCTC 1181
841 TCGGATCTGTTTTCTGCTCTCTCAAGCTGTTCACTTCTGCTCCGCTCCGACATCAAGC 900
1182 TCGGATCTGTTTTCTGCTCTCTCCCGAGCTGTTCACTTCTGCTCCGCTCCGACAGACA 1241
901 GTCAGGATGCAATGTTTCAATCTATCCCGGCGAGTACAGTCAACCGCATGCTTGG 960
1242 GTCAGGATGCAATGTTTCAATCTATCCCGGCGAGTACAGTCAACCGCATGCTTGG 1301
961 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
1302 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1361
1021 CCAACAAGCTGTGTGTCAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
1362 CCAACAAGCTGT 1421
1081 TACTATTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
1422 TACTATTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1481
1141 GTTGAACGGGGAACCTTACACAGACAGAGGGGGAACAAGCGCGCGCGCCACAGGGCTTACA 1200
1482 GTTGAACGGGGAACCTTACAGAGAGGGGGAACAAGCGCGCGCGCCACAGGGCTTACG 1541
1201 TCCCTCTTCAACCTTGGGCGGCTCAAAAATTCAGCTTGTAAACCAACAGGCACTGG 1260
1542 TCCCTCTTCAACCTTGGGCGGCTCAAAAATTCAGCTTGTAAACCAACAGGCACTGG 1601
1261 CACATCAACAGAACTGCTTGAACCTGCAATGATCTCCCTCCAACTGGGTTCTTGGCGGG 1320
1602 CACATCAACAGAACTGCTGCAACTGCAATGATCTCCCTCCAACTGGGTTCTTGGCTGG 1661

1321 CTGTTTACACGACAGGTTCAATGCTCCGAGATGCTTCAAGCGCATGGCCAGCTGCCCC 1380
1662 CTGTTTACACGACAGGTTCAATGCTCCGAGATGCTTCAAGCGCATGGCCAGCTGCCAGC 1721
1381 CCCATTTGACAGTTTCATCAGGGGTGGGTCCTCATCTTATATGATGCCAGGGCTTG 1440
1722 CCCATTTGACAGTTTCATCAGGGGTGGGTCCTCATCTTATATGATGCCAGGGCTTG 1781
1441 GACCAAGAGCCCTTATGCTGGGCACTACGACCTCAACCGTGTGTATGCTGCGCGTTG 1500
1782 GACCAAGAGCCCTTATGCTGGGCACTACGACCTCCGCGCGTGTATGCTACCCGCGCG 1841
1501 CAGGTGTGTGCGCCAGTGTACTTTCATCTCAAGCCCTGTTGTGTGGGAGCAAGCAT 1560
1842 CAGGTGTGTGCGCCAGTGTACTTTCATCTCAAGCCCTGTTGTGTGGGAGCAAGCAT 1901
1561 CGTTTGGGCGCCCTTACGTAACAGATGGGGTGAAGATAGACGAGCGTGTCTTCAAC 1620
1902 CGGTTCGCGCGTCTTACGTAACAGTGGGGGAGATAGACGAGCGTGTCTTCAAC 1961
1621 AACACGGGCGCCAGCGGCAACTGTTCCGCTGTACATGATGATAGACCGGGTTC 1680
1962 AACACGGGCGCCAGCGGCAACTGTTCCGCTGTACATGATGATAGACCGGGTTC 2021
1681 ACCAAGACGTGTGGGGCGCCCGCTGCAACATCGGGGGGTGGGCAACACACTTTGATC 1740
2022 ACCAAGACGTGTGGGGCGCCCGCTGTAACATCGGGGGGATCGGCAATTAACCTTGACC 2081
1741 TGCCCCACGACCTCTTCCGGAAGCATCCGAGGCGCACTTACCACAATGCGTTTGGGG 1800
2082 TGCCCCACGACCTCTTCCGGAAGCATCCGAGGCGCACTTACCACAATGCGTTTGGGG 2141
1801 CTTGGTTG 1809
2142 CTTGGTTG 2150

RESULT 14
AB008441
LOCUS
DEFINITION Hepatitis C virus RNA for polyprotein, partial cds, clone:15-2.
ACCESSION AB008441
VERSION AB008441.1 GI:3523046
KEYWORDS polyprotein.
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
Virus; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
REFERENCE
1 (sites)
AUTHORS Kato,N., Ikeda,M., Sugiyama,K., Mizutani,T., Tanaka,T. and Shimotohno,K.
TITLE Hepatitis C virus population dynamics in human lymphocytes and hepatocytes infected in vitro
JOURNAL U. Gen. Virol. 79 (Pt 8), 1859-1869 (1998)
MEDLINE 98378035
PUBMED 9714233
REFERENCE
2 (bases 1 to 3296)
AUTHORS Kato,N.
TITLE Direct Submission
JOURNAL Submitted (18-OCT-1997) Nobuyuki Kato, National Cancer Center Research Institute, Virology Division; 5-1-1 Tsukiji, Chuo-Ku, Tokyo 104, Japan (E-mail:nkato@ncr.90.jp, Tel:03-3542-2511(ex.4701), Fax:03-3543-2181)
FEATURES
location/Qualifiers
1..3296
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/isolate="1B-2"
/db_xref="taxon:11103"
/clone="15-2"
/cell_line="Hep-2C"
/cell_type="T-cell"
273..3296

CDS

```
/codon_start=1
/product="polypeptide"
/protein_id="BAJ32647.1"
/db_xref="GI:3523047"
/translation="MSTNPKPQKTKRNTNRPRQDYKFRGGQIVGGVYLPRGRRL
GVATNRTSRQPRGRQPIPRARRPERAAMAQPGYPMPLVGNBELGAGWMLISRG
SRPSMPTDPRRRSRNLGKVIDLTCGFDLMGYILVPAALPGASALAHGRVLED
SNVATGNLPGCSFSILFLASCLTI PASAYEVNVSQVYHVTNCSNLSIYEAD
MNMTHGCVPCVREANSRQWALPTLALNNSVPTTTIRHVDLVLGAAALCSMY
VGDIGSVFLISQLPRESRQWETVODCKSLSPHVSHTIRHMDMMMDMMALCLMY
SOLLIRPAVDPMVAGHWGLAGLAYSINVGMAVYLVMILPAVDGATTTSGCTV
GTTTGLTSFAPBQKLIQINTNSMHI NRTALCNDSLHGFALPIYHKFNSS
GSEBMSCRPIDKFAQGWGPI THGVDDPDORPCMYHAPRCGLVPAISQVGRVYC
FTSPVVGTTDRSGAPTYTNGENETDVLNNTRPQNMFGCTVNSNLSYFTKCG
PCNIGVGNNTLTCPTDCPRKREATVYKSGPMPLTRCWDVYRSLMYLPTCNF
TIFKRMVYGVGEBHLDACNMTREGRCVREDRSELSPLISTEMOVLPCSTPL
PALSTGLIHQNTIDVOYLIVGVSVSVKMEYVILFLILADARCACLMMMLL
VAQBALENLVLINAAVAGAGILSFVFECAAMYIKGKLVPGAAPVAYGWPILL
LILALPRAVYAMDREMAASCGGAVFGLALTLSPHYKFLARLIMWLYLTLRAL
LQWMLPELVNRRGDAIILLTCVHVELIFDIKILALFGLPLMVLQAGLTRVPEVR
AGLIRVCMVLRKVGHYVOMALMBLALITGYVYDHLTPLRGMAHAGRLDAVAE
PVFSDMETKIITMGADTAACGDIILGLPVSARRR"
```

ORIGIN

```
Query Match      86.8%; Score 1570.6; DB 14; Length 3296;
Best Local Similarity 91.8%; Pred. No. 0;
Matches 1660; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

QY 1 ATGAGCAGAAATCTTAAACCTCAAGAAAAAACAAGTAACCAACCGCGGCCACAG 60
DB 273 ATGAGCAGAAATCTTAAACCTCAAGAAAAAACAAGTAACCAACCGCGGCCACAG 332
QY 61 GACGTAAGTCCCGGCGGTGTGCTAGTCTGTGGTGAAGTTAAGTGGTCCGCGCAG 120
DB 333 GACGTAAGTCCCGGCGGTGTGCTAGTCTGTGGTGAAGTTAAGTGGTCCGCGCAG 392
QY 121 GGCCTCAGGTTGGGTGTGCGCGGAGTAAGAACTTCCGAGCGGTGCAACTCTGTGA 180
DB 393 GGCCTCAGGTTGGGTGTGCGCGGAGTAAGAACTTCCGAGCGGTGCAACTCTGTGA 452
QY 181 AGGCGACAACCTATCCCAAGGCTGCGCAGCGCCGAGGGGCTGGGCTCAAGCCCGG 240
DB 453 AGGCGACAACCTATCCCAAGGCTGCGCAGCGCCGAGGGGCTGGGCTCAAGCCCGG 512
QY 241 TACCTTGGCCCTCTATGAGCAACGAGGAGATGGGGTGGGCAAGATGCTCTGTACCC 300
DB 513 TACCTTGGCCCTCTATGAGCAATGAGGCTTGGGGTGGGCAAGATGCTCTGTACCC 572
QY 301 CGTGGCTCCGCGCTAGTGGGCGCCCACTGACCCCGGCGTAGTCCGTAATTGGGT 360
DB 573 CGTGGCTCCGCGCTAGTGGGCGCCCACTGACCCCGGCGTAGTCCGTAATTGGGT 632
QY 361 AAAGTCATGATACCTCTCAATGCGGCTTGCAGCCTCATGGGAGTACATTTCCGCTCTG 420
DB 633 AAAGTCATGATACCTCTCAATGCGGCTTGCAGCCTCATGGGAGTACATTTCCGCTCTG 692
QY 421 GCGCGCTCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTTCTGAGAGAC 480
DB 693 GCGCGCTCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTTCTGAGAGAC 752
QY 481 GCGGTGAACATGACAGGGAATTTAACCGGTTGCTTCTCTATCTTCTCTCTGGGCT 540
DB 753 GCGGTGAACATGACAGGGAATTTGCGCGGTTGCTTCTCTATCTTCTCTCTGGGCT 812
QY 541 TTGCTGTCTGTTTGAACATTCAGGCTTCCGTTATGAAATGGCAACGTTGCCGGATC 600
DB 813 TTGCTGTCTGTTTGAACATTCAGGCTTCCGTTATGAAATGGCAACGTTGCCGGATC 872
QY 601 TACCATGTACGAACGATTTGCTCAACTCAAGCATGTGTACGAGACAGGACATGATC 660
DB 873 TACCATGTACGAACGATTTGCTCAACTCAAGCATGTGTATGAGGACAGGACATGATC 932
QY 661 ATGCAACACCCCGGGGTGTGCTGTGCTCGGAGGGGTAAATTCCTCCGCTCTGGGTA 720
```

```
DB 933 ATGCAACACCCCGGGGTGTGCTGTGCTCGGAGGGCCCAACTCTCCGCTGTGGGTA 992
QY 721 GGGCTACCTCCACAGCTGTGGGGCCCAAGGACGGCAGCATCCCACTGCGCAATATGACAGC 780
DB 993 GGGCTACCTCCACAGCTGTGGGGCCCAAGGACGGCAGCATCCCACTGCGCAATATGACAGC 1052
QY 781 CACGTGATTTGCTCGTTGGGGCGGCTGCTTCTGCTCCGCTATGATGACGTGGGGATCTC 840
DB 1053 CACGTGATTTGCTCGTTGGGGCGGCTGCTTCTGCTCCGCTATGATGACGTGGGGATCTC 1112
QY 841 TGGGATCTGTTTCTCTGCTCTCTCAAGCTGTTCACTTCTCGCTTCCGCAATCAAGC 900
DB 1113 TGGGATCTGTTTCTCTGCTCTCTCAAGCTGTTCACTTCTCGCTTCCGCAATCAAGC 1172
QY 901 GTAACGAGTGAATTTGTTCAATCTATCCCGGCAAGTACAGTACCGGATGCTTGG 960
DB 1173 GTAACGAGTGAATTTGTTCAATCTATCCCGGCAAGTACAGTACCGGATGCTTGG 1232
QY 961 GATATGATGATGAATGATGATCACTACAGAGCCCTAGTGGTATGCGAGCTACTCGGATC 1020
DB 1233 GATATGATGATGAATGATGATCACTACAGAGCCCTAGTGGTATGCGAGTGTCTCGGATC 1292
QY 1021 CCAAGCTGTGTGTGACATGTGTGGGGGGCCCACTGGGAGTCTGGCGGCTTGGC 1080
DB 1293 CCAAGCTGTGTGTGACATGTGTGGGGGGCCCACTGGGAGTCTTACGCGGCTTGGC 1352
QY 1081 TACTATTCAATGTGTGGGAACTGGGCTAAGTCTTGTGTGATGATGATCTTTGGCGGC 1140
DB 1353 TACTATTCAATGTGTGGGAACTGGGCTAAGTCTTGTGTGATGATGATCTTTGGCGGC 1412
QY 1141 GTTGAAGGGGAACCTTACAGAGCAGAGGGGGAACAACGCGCGCGCCACAGGGCTTACA 1200
DB 1413 GTTGAAGGGGCTACCTACAGAGCAGAGGGGGAAGGATGCGGACACAGAGGCTTACG 1472
QY 1201 TCCCTCTTCAACCTTGGGCGGCTCGAAGAAATTCAGCTTGTAAACCAACGCGAGCTGG 1260
DB 1473 TCCCTCTTCAACCTTGGGCGGCTCGAAGAAATTCAGCTTGTAAACCAACGCGAGCTGG 1532
QY 1261 CACATTAACAGAACTGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
DB 1533 CACATTAACAGAACTGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1592
QY 1321 CTGTTTCAACGAGAGTTCATATGCGTCCGATGCTGAGAGGACATGAGGCTTGGTGG 1380
DB 1593 CTGTTTCAACGAGAGTTCATATGCGTCCGATGCTGAGAGGACATGAGGCTTGGTGG 1652
QY 1381 CCCATTGACCAAGTTCGATCAGGAGGTGGGATCCCATCACTTATATGATGCCACGAGCTTG 1440
DB 1653 CCCATTGACCAAGTTCGATCAGGAGGTGGGATCCCATCACTTATATGATGCCACGAGCTTG 1712
QY 1441 GACAGAGGCTCTATGCTGCGCATACGCACTCAACCGTGTGTATGCTGCGCGGTTG 1500
DB 1713 GACAGAGGCTCTATGCTGCGCATACGCACTCAACCGTGTGTATGCTGCGCGGTTG 1772
QY 1501 CAGGTGTGTGCGCAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
DB 1773 CAGGTGTGTGCGCAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1832
QY 1561 CGTTTGGCGCCCTTACATGATGAGTGGGTGAGATGAGACGATGCTGCTTCTCAAC 1620
DB 1833 CGTTTGGCGCCCTTACATGATGAGTGGGTGAGATGAGACGATGCTGCTTCTCAAC 1892
QY 1621 AACACGCGCGCCGACAGGAGCACTGCTTGGCTGTATGATGATGATGATGATGATGATGAT 1680
DB 1893 AACACGCGCGCCGACAGGAGCACTGCTTGGCTGTATGATGATGATGATGATGATGATGAT 1952
QY 1681 ACCAAGAGTGTGGGGGCGCCCGTGTGCAACATGCGGGGGGTGGGCAACAACCTTGTATC 1740
DB 1953 ACCAAGAGTGTGGGGGCGCCCGTGTGCAACATGCGGGGGGTGGGCAACAACCTTGTATC 2012
QY 1741 TGGCCACAGGACTGCTTCCGGAAGATCCGAGGCGCATTAACCAAAATGCGGTTGGGG 1800
```


Db	2013	TGCCCCAGGACTGCTTCCGGAAGACACCCGAGCCACTTACACCAATATGCGGCTCGGG	2012
Oy	1801	CCTTGTTG 1809 	
Db	2073	CCATGGTTG 2081	
RESULT 15			
LOCUS	HCJ238800	9033 bp	linear VRL 18-AUG-1999
DEFINITION	Hepatitis C virus type 1b complete genome, isolate NC1.		
ACCESSION	AJ238800		
VERSION	AJ238800.1	GI:5748510	
KEYWORDS	complete genome; core protein; glycoprotein E1; glycoprotein E2; NS2 proteinae; NS3 proteinase/helicase; NS3/4A proteinase cofactor; NS4b protein; NS5A phosphoprotein; NS5B RNA dependant RNA polymerase; p7 peptide; polypotein.		
SOURCE	Hepatitis C virus type 1b		
ORGANISM	Hepatitis C virus type 1b		
	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.		
REFERENCE	1		
AUTHORS	Koch J.O. and Bartenschlager R.		
TITLE	Mutation of hepatitis C virus NS5A hyperphosphorylation by nonstructural proteins NS3, NS4A, and NS4B		
JOURNAL	J. Virol. 73 (9), 7138-7146 (1999)		
MEDLINE	99370154		
PUBMED	10438800		
REFERENCE	2 (bases 1 to 9033)		
AUTHORS	Bartenschlager, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (10-MAY-1999) Bartenschlager R., Institute for Virology, Johannes Gutenberg - University Mainz, Obere Zahnbacher Strasse 67, 55131 Mainz, GERMANY		
FEATURES	Location/Qualifiers		
source	1..9033		
	/organism="Hepatitis C virus type 1b"		
	/villon		
	/mol_type="genomic RNA"		
	/strain="RB"		
	/isolate="NC1"		
	/db_xref="taxon:31647"		
	/country="Germany"		
	1..9033		
	/codon_start=1		
	/product="polyprotein"		
	/protein_id="CAB51095.1"		
	/db_xref="GI:5748511"		
	/db_xref="GDA:O9Q0P6"		
	/db_xref="UniProt:K1REMBL:O9Q0P6"		
	/translation="MSTNPKPQRKTKNTNRPODVKFPGGQIVGGVILLPRGPRLTGVRATKTPRSRPSOPRGPROPIEKARQEPGRWADQGVWPLVGNBGIQMGWGLISPPGSRPMSGTDRPRRRNRGKVIDTLTCEPADLMGTLVGLVAPGLGAAALAHGVRVLDGVRATGNGPCGSFIFLALISCLTTPASAYEVNRNVSIVHNDSCNASIYEADMIMHTPGCVPCVRNNSSRCWVALPTPLAANASVPTTTRRHVDLLVGAALCSANYVGDLCGSVFVLAQLFTFSPRHRHETVODNCISITYGHVTHGRHMDMMNNSPIALAVSOLLRIPOAVVDVAVAGAHMGVLAGAYSMGNMAKYLIVMLPAGVDGGVYVGTMMAKNTLITLSFSPGSSOKLOIYNNGNSMHNRTLNCDLSINTGFLAIVYHKENSSCCPERMASCTIDAPAQMGPTITYNHSSDQRYCHVAPRPGCIVPAQVCGPVYCTPSPVAVVGTITRFGVPTPYSNGENETDVLNNTNRPQGMWFGCTWNSVGFPTKCGGFPNNGIGSGKTLCTPTDCEPKHPEATYTKGSDPMLIPCLVHYPYRLAHYPCVWFITIFKRVNVGVEHRELEACNMTGGERCNLEDRSGDPLISPLISTETMOVLPCSPFTLIPALSTGILIHQNVVDVOYLIGISAVSFAIKREYVILLILLILADARVCAIMMLLILAOEATLENTVYNAASVAGAHGISLSPVFCAMTITKRLVNGAAYALXGWPILLILALPPOAVAMDEMASCGGAVFVGLILTTSFPHKFEVTRPTMWLYFTTAEAHLQWVIPLNVAGSDAVILLTCAIHPELIFTITLILALIGLIMVLAQGITKYVPRAPHGIRACMLVRYKAGVHYVOMALMLKLAALGTIVYHILPLRDMAHAGRDIAVAVEVSDMETKVTITWGPDPAAAGDILIGLPVBARGRREIHGPNLSIEGOGMRLIAHTAYSOOTGILGCIITSLTGRDRNOVEGVWVGNATROSLATCNGVCGVYHAGSKTLAPKKEPTIOTMTYNNDDLVGMQAPRPAARLTPTCTGSSDLYIVTTHAVITPYRRRGDTRSKLLSPRVSYLTKSSGGPLLCPESHAVGITRAAVVCTRGAKADVPEVSMETTNRSPVETDNNSPPAVPQTFQVAHIAHPGSGSKTKVPAAVAAQYKTVLVPNSVAULTSPGVAHNSAGHIDENIRGTGVTITTTGAAPIYSTVTKGLADGGCGGAYDILICDECHS		

[illegible]


```

ORIGIN
Query Match      86.8%; Score 1570.6; DB 14; Length 9033;
Best Local Similarity 91.8%; Pred. No. 0;
Matches 1660; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

gene      /gene="NSSA"
          /product="NSSA phosphoprotein"
          /size="9030"
          /gene="NSSB"
          /size="9030"
          /gene="NSSB"
          /product="NSSB RNA dependant RNA polymerase"

mat_peptide

QY 1 ATGAGCAGCAATCTCTTAACCTCAAGAAAAACCAACGTAACCAACCGCCGCCACAG 60
Db 1 ATGAGCAGCAATCTCTTAACCTCAAGAAAAACCAACGTAACCAACCGCCGCCACAG 60
QY 61 GACGTCAAGTCCCGGCGGAGTGTAGATCGTTGTGTGAGATTACCTGTTGCCGCGAG 120
Db 61 GACGTCAAGTCCCGGCGGAGTGTAGATCGTTGTGTGAGATTACCTGTTGCCGCGAG 120
QY 121 GGCCTCAGGTTGGGTGTGCGCGGAGTAAAGAGACTTCGAGCGGTGCAACCTGCTGA 180
Db 121 GGCCTCAGGTTGGGTGTGCGCGGAGTAAAGAGACTTCGAGCGGTGCAACCTGCTGA 180
QY 121 GGCCTCAGGTTGGGTGTGCGCGGAGTAAAGAGACTTCGAGCGGTGCAACCTGCTGA 180
Db 121 GGCCTCAGGTTGGGTGTGCGCGGAGTAAAGAGACTTCGAGCGGTGCAACCTGCTGA 180
QY 181 AGGCGCAACCTATCCCAAGGCTCGCCAGCCGAGGCGAGGCGCTGGGCTCAAGCCGAG 240
Db 181 AGGCGCAACCTATCCCAAGGCTCGCCAGCCGAGGCGAGGCGCTGGGCTCAAGCCGAG 240
QY 241 TACCTTGGAGCCCTCTATGAGCAACAGAGGAGTGGGAGAGATGAGCTCTGTACCC 300
Db 241 TACCTTGGAGCCCTCTATGAGCAACAGAGGAGTGGGAGAGATGAGCTCTGTACCC 300
QY 301 CGTGGCTCCGCGCTAGTGGGAGCCCACTGACCCCGCGGAGTGTGCGTAATTTGGGT 360
Db 301 CGTGGCTCCGCGCTAGTGGGAGCCCACTGACCCCGCGGAGTGTGCGTAATTTGGGT 360
QY 361 AAAAGTATGATATCCCTCAATGCGGCTTGCAGCACTCAATGGGATCAATTCGGCTGTC 420
Db 361 AAAAGTATGATATCCCTCAATGCGGCTTGCAGCACTCAATGGGATCAATTCGGCTGTC 420
QY 421 GGCCTCAGGTTGGGTGTGCGCGGAGTAAAGAGACTTCGAGCGGTGCAACCTGCTGA 480
Db 421 GGCCTCAGGTTGGGTGTGCGCGGAGTAAAGAGACTTCGAGCGGTGCAACCTGCTGA 480
QY 481 GGCCTCAGGTTGGGTGTGCGCGGAGTAAAGAGACTTCGAGCGGTGCAACCTGCTGA 540
Db 481 GGCCTCAGGTTGGGTGTGCGCGGAGTAAAGAGACTTCGAGCGGTGCAACCTGCTGA 540
QY 541 TTGCTGTCTGTGTGACATTCAGCTTCGCTTAAGAGTGGCAACGCTGTCCGAGATC 600
Db 541 TTGCTGTCTGTGTGACATTCAGCTTCGCTTAAGAGTGGCAACGCTGTCCGAGATC 600
QY 601 TACCATGTCAACGATGCTCCAACTCAAGCATGCTGTAAGAGACAGCGGACATGATC 660
Db 601 TACCATGTCAACGATGCTCCAACTCAAGCATGCTGTAAGAGACAGCGGACATGATC 660
QY 661 ATGCAACCCCGGAGTGTGCGCTGTGTCGAGGAGGTAATTCCTCCGCTGCTGGAT 720
Db 661 ATGCAACCCCGGAGTGTGCGCTGTGTCGAGGAGGTAATTCCTCCGCTGCTGGAT 720
QY 721 GGCCTCAGGTTGGGTGTGCGCGGAGTAAAGAGACTTCGAGCGGTGCAACCTGCTGA 780
Db 721 GGCCTCAGGTTGGGTGTGCGCGGAGTAAAGAGACTTCGAGCGGTGCAACCTGCTGA 780
QY 781 CACGTGATTTGCTGTGAGGAGGAGTGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db 781 CACGTGATTTGCTGTGAGGAGGAGTGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 840
QY 841 TGGGATCTGTTTCTCTGCTCTCTCAAGCTTCACTTCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 841 TGGGATCTGTTTCTCTGCTCTCTCAAGCTTCACTTCTGCTGCTGCTGCTGCTGCTGCTG 900

```

```

QY 901 GTACAGAGCTGCAATTTGTTCAATCTATCCCGGCAAGTATCAGTCAACCGAGCTTGG 960
Db 901 GTACAGAGCTGCAATTTGTTCAATCTATCCCGGCAAGTATCAGTCAACCGAGCTTGG 960
QY 961 GATATGATGATGAATGCTGCTACCTCAAGAGCTTATGATATGAGCTACTCTCCGATC 1020
Db 961 GATATGATGATGAATGCTGCTACCTCAAGAGCTTATGATATGAGCTACTCTCCGATC 1020
QY 1021 CCACAGCTGTGTGAGCAATGCTGTGCGGAGGCGCCACTGGGAGATCCTGGCGGCTTGGC 1080
Db 1021 CCACAGCTGTGTGAGCAATGCTGTGCGGAGGCGCCACTGGGAGATCCTGGCGGCTTGGC 1080
QY 1081 TACTATTCATGATGAGGAACTGGGCTAAGTCTTGTTGATGATGCTACTTTTGGCGGC 1140
Db 1081 TACTATTCATGATGAGGAACTGGGCTAAGTCTTGTTGATGATGCTACTTTTGGCGGC 1140
QY 1141 GTTGAAGGAGAACTTACAGCAAGAGGAGAGCAACGAGCGGCGCCACGAGCTTACA 1200
Db 1141 GTTGAAGGAGAACTTACAGCAAGAGGAGAGCAACGAGCGGCGCCACGAGCTTACA 1200
QY 1201 TCCCTTTTCAACCTGAGGCGGCTCAAGAAATTCAGCTTGTAAACCAACGAGAGCTGG 1260
Db 1201 TCCCTTTTCAACCTGAGGCGGCTCAAGAAATTCAGCTTGTAAACCAACGAGAGCTGG 1260
QY 1261 CACATCAACAGAACTGCTTGAACCTGCAATGACTCCCTCAAACTGGTTCCTTGGCGCG 1320
Db 1261 CACATCAACAGAACTGCTTGAACCTGCAATGACTCCCTCAAACTGGTTCCTTGGCGCG 1320
QY 1321 CTGTTTCAACGACAGTTCATATGCTGCGAGTGTCAAGGAGAGAGAGAGAGAGAGAGAG 1380
Db 1321 CTGTTTCAACGACAGTTCATATGCTGCGAGTGTCAAGGAGAGAGAGAGAGAGAGAGAG 1380
QY 1381 CCGATTGACAGTTCATATGAGGAGTGGGCTCCCATCTATTAATAGTCCACGAGCTTG 1440
Db 1381 CCGATTGACAGTTCATATGAGGAGTGGGCTCCCATCTATTAATAGTCCACGAGCTTG 1440
QY 1441 GACCAAGAGCCCTATTTGCTGAGCACTAGCACTCAACCGTGTGATGCTGCGCGGCTTG 1500
Db 1441 GACCAAGAGCCCTATTTGCTGAGCACTAGCACTCAACCGTGTGATGCTGCGCGGCTTG 1500
QY 1501 CAGGTGTGATGCTCAAGTGTACTGCTTCACTCAAGCCCTGTTGGTGGGAGAGAGAGAG 1560
Db 1501 CAGGTGTGATGCTCAAGTGTACTGCTTCACTCAAGCCCTGTTGGTGGGAGAGAGAGAG 1560
QY 1561 CCGTTTGGGCGCCCTTCACTGATGAGTGGGAGTGAATGAACAGGACCTGCTGCTTCAAC 1620
Db 1561 CCGTTTGGGCGCCCTTCACTGATGAGTGGGAGTGAATGAACAGGACCTGCTGCTTCAAC 1620
QY 1621 AACAGCGGCGGCAAGGAGCAACTGCTGAGCTGTACATGATGATGATGATGATGATGATG 1680
Db 1621 AACAGCGGCGGCAAGGAGCAACTGCTGAGCTGTACATGATGATGATGATGATGATGATG 1680
QY 1681 ACCAAGCTGTGAGGAGCCCGGCTGCAACATCGAGGAGGATCGGCAACAACTTTGATC 1740
Db 1681 ACCAAGCTGTGAGGAGCCCGGCTGCAACATCGAGGAGGATCGGCAACAACTTTGATC 1740
QY 1741 TGGCCCAAGGAGCTGCTTCCGGAAGCATCCCGAGGCGCACTTACACCAAAATGCGGTTGG 1800
Db 1741 TGGCCCAAGGAGCTGCTTCCGGAAGCATCCCGAGGCGCACTTACACCAAAATGCGGTTGG 1800
QY 1801 CTTTGGTTG 1809
Db 1801 CTTTGGTTG 1809

```

Search completed: February 19, 2005, 07:08:48
Job time : 8047.09 secs

THIS PAGE BLANK (USPTO)

CC This sequence is thought to encode viral structural and non-structural
CC proteins of the PT-NANBH viral genome which are antigenic. It was
CC isolated from human serum infections for the virus. See also AAQ12236-41.
CC (updated on 25-MAR-2003 to correct PA field.)

SQ Sequence 2116 BP; 392 A; 650 C; 624 G; 450 T; 0 U; 0 Other;

Query Match	99.8%	Score 1805.8	DB 2	Length 2116
Best Local Similarity	99.9%	Pred. No. 0		
Matches 1807; Conservative	0	Mismatches 2	Indels 0	Gaps 0

QY	1	ATGACAGAGATCCCTAAACCTCAAGAAAAACCAACCTAACACCAACCGCGCCCAAG	60
Db	308	ATGAGCAAGAAATCCTAAACCTCAAGAAAAACCAAGCAACCAACCGCGCCCAAG	367
QY	61	GACGTCAAGTTCCCGGGGCGTGGTCAATCGTTGGTGAAGTTTACTTTCGCGCGCAG	120
Db	368	GACGTCAAGTTCCCGGGGCGTGGTCAATCGTTGGTGAAGTTTACTTTCGCGCGCAG	427
QY	121	GGCCCAAGTTGGGTGGCGGCGCACTAGGAAGACTTCCAGCGGTCCCAACTTGTGGA	180
Db	428	GGCCCAAGTTGGGTGGCGGCGCACTAGGAAGACTTCCAGCGGTCCCAACTTGTGGA	487
QY	181	AGGCGACAACCTATCCCAAGGCTGCGCAGGCCCGAGGGGCGCTGGGGCTCAAGCCGG	240
Db	488	AGGCGACAACCTATCCCAAGGCTGCGCAGGCCCGAGGGGCGCTGGGGCTCAAGCCGG	547
QY	241	TACCTTGGCCCTCTTATAGGACAGAGGGACATGGGGTGGGCGAGATGAGCTCCTGTCA	300
Db	548	TACCTTGGCCCTCTTATAGGACAGAGGGACATGGGGTGGGCGAGATGAGCTCCTGTCA	607
QY	301	CGTGGCTCCCGGCGTAAAGTTGGGGGCCCCACTGACCCCCGGGCGTAAAGTCCGCTA	360
Db	608	CGTGGCTCCCGGCGTAAAGTTGGGGGCCCCACTGACCCCCGGGCGTAAAGTCCGCTA	667
QY	361	AAAGTCATGATACCTCTACATAGCGGCTTCGCGCAGCTCATAGGGGTACATTCCGCTCG	420
Db	668	AAAGTCATGATACCTCTACATAGCGGCTTCGCGCAGCTCATAGGGGTACATTCCGCTCG	727
QY	421	GCGCTCCCTTAAGGGGGCGCTGCCAGGGGCTTGGCGCATGGCGTCCGGGTTCTGAGGAC	480
Db	728	GCGCTCCCTTAAGGGGGCGCTGCCAGGGGCTTGGCGCATGGCGTCCGGGTTCTGAGGAC	787
QY	481	GGCGTGAATATAGCAACAGGGGAATTACCGGGTGTCTTTCTCATCTTCTCTGTGGCT	540
Db	788	GGCGTGAATATAGCAACAGGGGAATTACCGGGTGTCTTTCTCATCTTCTCTGTGGCT	847
QY	541	TTGCTGTCTGTTTGAACCATTCAGAGCTTCGGTTATGAAGTGGCAGCGTTCGCGGATC	600
Db	848	TTGCTGTCTGTTTGAACCATTCAGAGCTTCGGTTATGAAGTGGCAGCGTTCGCGGATC	907
QY	601	TACCATATGACGAGACGATTTGCTCCACATCAGCATCGTATAGAGACAGCGGACATGATC	660
Db	908	TACCATATGACGAGACGATTTGCTCCACATCAGCATCGTATAGAGACAGCGGACATGATC	967
QY	661	ATGCAACACCCCGGGGTGTGTGCCGTGTCGCGAGGGGTAAATCTCCCGGCTGGGGTA	720
Db	968	ATGCAACACCCCGGGGTGTGTGCCGTGTCGCGAGGGGTAAATCTCCCGGCTGGGGTA	1027
QY	721	GCGCTCACTCCACGCTCGCGGCCAAGACGCGACAGATCCCACTGCGACAATAGAGCG	780
Db	1028	GCGCTCACTCCACGCTCGCGGCCAAGACGCGACAGATCCCACTGCGACAATAGAGCG	1087
QY	781	CACGTGCAATTTGCTGTGGGGGCGGCTGCTTGTCTCCGTATATACGTGGGGATATTC	840
Db	1088	CACGTGCAATTTGCTGTGGGGGCGGCTGCTTGTCTCCGTATATACGTGGGGATATTC	1147
QY	841	TGCGGATGTGTTTCTCGTCTCTAGCGTTCACCTTCGCGCTCGCGACATCAGAG	900
Db	1148	TGCGGATGTGTTTCTCGTCTCTAGCGTTCACCTTCGCGCTCGCGACATCAGAG	1207
QY	901	GTACAGAGTGCATTTGTTCAATCTATCCGGGCAAGTATCAGTACCGCATGGCTTGG	960

Db	1208	GTACAGAGCTGCATTGTTCAATCTATTCGCCGCCACGTATCAGGTTCACCGCATGCTTGG	1267
Qy	961	GATATGATGATGAACCTGTGTCACCTACAGACAGCCCTAGTGTATGACGCTATCTCCGAGTC	1020
Db	1268	GATATGATGATGAACCTGGTCACTACAGACAGCCCTAGTGTATGACGCTATCTCCGAGTC	1327
Qy	1021	CCACAAGCTGTCTGTGACATATGTTGGCGGGGCCCACTGGGGAGTCTTGGCGGGCTTGGCC	1080
Db	1328	CCACAAGCTGTCTGTGACATATGTTGGCGGGGCCCACTGGGGAGTCTTGGCGGGCTTGGCC	1387
Qy	1081	TACTATTTCATGATGTGGGGAACTGGGAGTAAAGGCTTGTGGTGTGATGCTACTTTGGCCGGC	1140
Db	1388	TACTATTTCATGATGTGGGGAACTGGGAGTAAAGGCTTGTGGTGTGATGCTACTTTGGCCGGC	1447
Qy	1141	GTTGACGGGGAACTTTACACGACAGGGGGGACACACGGCGCGCCGCCCAAGGCTTTACA	1200
Db	1448	GTTGACGGGGAACTTTACACGACAGGGGGGACACACGGCGCGCCGCCCAAGGCTTTACA	1507
Qy	1201	TCCCTCTTCAACACTGGGGCGGGCTCAAGAAATCCAGCTTTGTAACACCAACGGCAGCTGG	1260
Db	1508	TCCCTCTTCAACACTGGGGCGGGCTCAAGAAATCCAGCTTTGTAACACCAACGGCAGCTGG	1567
Qy	1261	CACATTCACACAACTGCGCTTGAATGTGAATGACCTCCCTCCAAATCGGGTTCCTTGGCCGC	1320
Db	1568	CACATTCACACAACTGCGCTTGAATGTGAATGACCTCCCTCCAAATCGGGTTCCTTGGCCGC	1627
Qy	1321	CTGTTTTCACACGACAGGTTCAATGCGTCCGGATGCTCAGAGCGCATGCGCAGCTGCGC	1380
Db	1628	CTGTTTTCACACGACAGGTTCAATGCGTCCGGATGCTCAGAGCGCATGCGCAGCTGCGC	1687
Qy	1381	CCCATTCACACAGTTGCGATCAGGGGTGGGGTCCCATCACTTAATGATGCCACGCGCTTG	1440
Db	1688	CCCATTCACACAGTTGCGATCAGGGGTGGGGTCCCATCACTTAATGATGCCACGCGCGTTG	1747
Qy	1441	GACCAAGGGCCCTTAATGCTGGCACACTAGCACCTCAACCGGTGGTATCGTGGCCGGCGTTG	1500
Db	1748	GACCAAGGGCCCTTAATGCTGGCACACTAGCACCTCAACCGGTGGTATCGTGGCCGGCGTTG	1807
Qy	1501	CAGGTGTGTGGCCCAAGTACTGTTTCACTTCAAGCCCTGTTGTGTGGGACGACCCGAT	1560
Db	1808	CAGGTGTGTGGCCCAAGTACTGTTTCACTTCAAGCCCTGTTGTGTGGGACGACCCGAT	1867
Qy	1561	CGTTTCGGCGCCCTTACGTACAGATGGGGGTGAGATGAGACGACGTCGTCTTCTCAAC	1620
Db	1868	CGTTTCGGCGCCCTTACGTACAGATGGGGGTGAGATGAGACGACGTCGTCTTCTCAAC	1927
Qy	1621	AACACGGCGCCCGCACGGGGGCACTTGGTTCGGCTGTACATGAGATGAAATAGCACCGGGTTC	1680
Db	1928	AACACGGCGCCCGCACGGGGGCACTTGGTTCGGCTGTACATGAGATGAAATAGCACCGGGTTC	1987
Qy	1681	ACCAAGACGTGTGGGGGCGCCCGCTGACAACTCGGGGGGGGTCCGCAACAACCTTTGATC	1740
Db	1988	ACCAAGACGTGTGGGGGCGCCCGCTGACAACTCGGGGGGGGTCCGCAACAACCTTTGATC	2047
Qy	1741	TGCCCCACGACCTGCTTCCGGAGACATCCGAGGCGCACTTACACCAATGCGGTTCCGGG	1800
Db	2048	TGCCCCACGACCTGCTTCCGGAGACATCCGAGGCGCACTTACACCAATGCGGTTCCGGG	2107
Qy	1801	CCTTGGTTG 1809	
Db	2108	CCTTGGTTG 2116	
RESULT 2			
ABK91431			
ID	ABK91431 standard; DNA; 9605 BP.		
XX	ABK91431;		
XX	15-NOV-2002 (first entry)		
XX	Hepatitis C virus Con 1 isolate DNA mutant 8.		
XX			

RESULT 2
ABK91431
ID ABK91431 standard; DNA; 9605 BP.

DT 15-NOV-2002 (first entry)

DE Hepatitis C virus Con 1 isolate DNA mutant 8.

OS	Hepatitis C virus.
XX	Synthetic.
FT	Key
FT	CDS
FT	Location/Qualifiers
FT	342..9374
FT	/tag= a
FT	/product= "HCV polypeptide"
FT	/note= "The polypeptide consists of the Core, E1, E2, P7,
FT	NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
FT	replace(6934,T)
FT	/tag= b
PN	WO200259321-A2.
XX	
PD	01-AUG-2002;
PF	16-JAN-2002; 2002MO-EP000526.
PR	23-JAN-2001; 2001US-0263479P.
XX	
PA	(RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
PI	De Francesco R, Migliaccio G, Paonessa G;
DR	WPI, 2002-599793/64.
XX	
PT	New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
PT	NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
PT	ribosome entry site (IRES) region, useful in studying HCV replication and
PT	expression.
XX	
PS	Claim 9; Page; 69pp; English.
XX	
CC	The invention relates to nucleic acid molecules comprising altered HCV
CC	NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
CC	internal ribosome entry site (IRES) region coding for one or more NS3,
CC	NS5A, or EMCV IRES mutations, respectively. The location of the mutations
CC	are detailed in the specification. Also included are (1) an expression
CC	vector comprising a nucleotide sequence coding for the altered nucleic
CC	acids, which is transcriptionally coupled to an exogenous promoter; (2) a
CC	recombinant cell human hepatoma cell comprising the altered nucleic acids
CC	; (3) a recombinant cell produced by introducing into a human hepatoma
CC	cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
CC	replicon enhanced cell or which containing a functional HCV replicon; (5)
CC	an HCV replicon enhanced cells made in the method; and (6) measuring the
CC	ability of a compound to affect HCV activity. The HCV replicons and HCV
CC	replicon enhanced cells are useful in studying HCV replication and
CC	expression, and HCV and host cell interactions, producing HCV RNA and
CC	proteins, and providing a system for measuring the ability of a compound
CC	to modulate one or more HCV activities e.g. to discover drugs which may
CC	treat HCV mediated diseases such as liver failure, cirrhosis and
CC	hepatocellular carcinoma. The present sequence is an HCV replicon Con 1
CC	mutant of the invention. Note: The present sequence is not shown in the
CC	specification but was created by the indexer using the HCV sequence
CC	appearing as ABK91411 and the information in Claim 9
XX	
SQ	Sequence 9605 BP; 1910 A; 2884 C; 2733 G; 2078 T; 0 U; 0 Other;
XX	
Query Match	86.9%; Score 1572.2; DB 6; Length 9605;
Best Local Similarity	91.8%; Pred.No. 0; Mismatches 148; Indels 0; Gaps 0;
Matches 161; Conservative	0;
DG	1 ATGACGACGAATCTTAAACTCAAGAAAAAACCAACGTAAACCAACCGCCGCACAG 60
DG	342 ATGACGACGAATCTTAAACTCAAGAAAAAACCAACGTAAACCAACCGCCGCACAG 401
DG	61 GACGTCGAATTGCCGGCGCGTGTCGATCTGTTGGAGATTAACTGTTGCCGCGCAG 120
DG	402 GACGTCGAATTGCCGGCGCGTGTCGATCTGTTGGAGATTAACTGTTGCCGCGCAG 461

QY	121	GGCCCCAGGTTGGGATGAGCGCGCACTAGAAAGACTTCGAGCGAGTCCGCAACTCTGATGA	180
Db	462	GGCCCCAGGTTGGGATGAGCGCGCACTAGAAAGACTTCGAGCGAGTCCGCAACTCTGATGA	521
QY	181	AGGCGACAACCTATCCCGAAGGCTCGCGAGCCGAGAGGACAGGACCTGAGGCTCAGCCCGGG	240
Db	522	AGGCGACAACCTATCCCGAAGGCTCGCGAGCCGAGAGGATAGGAGCTCGAGCCCGGG	581
QY	241	TACCTTGGCCCTCTATGCGACAGAGGACATGGAGTGGGATGGCTCTCTGTCAACC	300
Db	582	TACCCCTGGCCCTCTATGCGACAGAGGACTGGAGTGGGATGGCTCTCTGTCAACC	641
QY	301	CGTGGCTCCGGCCAGTTGGGGGCCCATGACCCCGGCGTAGGTGCGGAAATTTGGGT	360
Db	642	CGTGGCTTCGGCTTAGTTGGGGGCCCATGACCCCGGCGTAGGTGCGGAAATTTGGGT	701
QY	361	AAAGTCATGATACCCCTCACTCATGCGGCTTCGCGACCTCATGAGGATACATTCGCTGTC	420
Db	702	AAGGTCATGATACCCCTCACTCATGCGGCTTCGCGACCTCATGAGGATACATTCGCTGTC	761
QY	421	GGCGTCCCTTAGGGGGGCGTCCGACAGGACCTGGGCGATGGCCGTCCGGGTTCTGAGGAC	480
Db	762	GGCGCCCCCTTAGGGGGGCGTCCGACAGGACCTGGGCGATGGCCGTCTGAGGAGAC	821
QY	481	GGCGGAACTATGCAACAGGAAATTTACCCGGTGTCTTCTCTATTTCTCTGGCT	540
Db	822	GGCTGAACTATGCAACAGGAAATCTCGCGAGTGTCTTCTATTTCTTTGGCT	881
QY	541	TTGCTGCTCTTTGACCATTTCCAGCTTCCGCTTATGAATGCGCAACTGTCCGGAGTC	600
Db	882	TTGCTGCTCTTTGACCATTTCCAGCTTCCGCTTATGAATGCGCAACTGTCCGGAGTC	941
QY	601	TACCATGTCAGAAAGATTTGCTCCACTCAGATTCGATGACGAGACGGGACATGATC	660
Db	942	TACCATGTCAGAAAGACTGTCTCCACGCAAGACTTGATATGAGCAGCGGACATGATC	1001
QY	661	ATGCAACACCCCGGAGTGTGTCCTCGTGTCCGAGAGGGTAAATTTCTCCGCTGAGGTA	720
Db	1002	ATGCACTACCCCGGAGTGTGTCCTCGTGTGGGAGAACTCTCCGCTGTGCGGTA	1061
QY	721	GCGCTCACTCCACGCTGCGGCGCAAGACGTCAGACCTCCCACTGCGACATTAACAGCG	780
Db	1062	GCGCTCACTCCACGCTGCGGCGCAAGAACCTTAGCGTCCCACTACGACATTAACAGCG	1121
QY	781	CACGTGCAATTTGCTCTGTTGGGGCGGCTGCTTGTCTCCGCTATGTAAGTGGGGAATTC	840
Db	1122	CATGTCGATTTGCTCTGTTGGGGCGGCTGCTGCTCTCGCTATGTAAGTGGGGAATTC	1181
QY	841	TGCGGATCTGTTTTCTCGTCTCTCAGCTGTTCACCTTCGCGCTCGCGACATCAGAGG	900
Db	1182	TGCGGATCTGTTTTCTCGTCTCGCGCAAGCTGTTCACCTTCGCGCTCGCGGCAAGAGCA	1241
QY	901	GTACAGACTSCAATTTGTTCAATCTATCCGGGCAAGTATCAGGTCACCGCATGGCTTGG	960
Db	1242	GTACAGACTSCAATTTGTTCAATATCTCCGGGCAAGTATCAGGTCACCGTATGGCTTGG	1301
QY	961	GATATGATGATGAAGTGGTCACTACAGACACCTTAGTGGATATGGAGCTACTCGGATC	1022
Db	1302	GATATGATGATGAAGTGGTCACTACAGACACCTTAGTGGATATGGAGCTACTCGGATC	1361
QY	1021	CCACAAGCTGTCGATGACATGATGGCGGGGCCCACTAGGAGAGTCTTGGCGGCGCTTGGC	1080
Db	1362	CCACAAGCTGTCGATGATGATGGCGGGGCCCAATGGGGAAGTCTTAGCGGGCGCTTGGC	1422
QY	1081	TACTATTCATGATGGTGGGAACTGGGCTAAGGCTTTGGTGTGATGCTACTTTTGGCGGC	1140
Db	1422	TACTATTCATGATGGTGGGAACTGGGCTAAGGCTTTGATGTGATGCTACTTTTGGCGGC	1481
QY	1141	GTTGACGGGGAACTTATCAGACAGGGGGGACACACGGCGCGCGCCACAGGGCTTACA	1200
Db	1482	GTTGACGGGGAACTTATGTGACAGGGGGGACGATGGCCAAAACACCTCTGGGATTTACG	1541

```
QY 1201 TCCCTTTGACACTGGGCGGCTCGAAGAAATCCAGCTTTAAACCAACGAGCACTGG 1260
    |||||
XX 1542 TCCCTCTTTTCAACCGGGGTCATCCGAGAAATCCAGCTTTAAACCAACGAGCACTGG 1601
    |||||
QY 1261 CACATCAACAGAACTGCTTTGAATGCACTGACTCCCTCCAAATGCGGTTCTTTCGCGG 1320
    |||||
DB 1602 CACATCAACAGAACTGCTTTGAATGCACTGACTCCCTCCAAATGCGGTTCTTTCGCGG 1661
    |||||
QY 1321 CTGTTCTACGACGACAGTTTCATGCGTCCGAGTGTGAGAGCGAGCGAGCTGCGCG 1380
    |||||
DB 1662 CTGTTCTACGACGACAGTTTCATGCGTCCGAGTGTGAGAGCGAGCGAGCTGCGCG 1721
    |||||
QY 1381 CCCATTGACCACTTCATGACGAGGAGGAGGAGTCCCATCACTTAATATAGTCCGAGCGTTG 1440
    |||||
DB 1722 CCCATTGACGAGGAGTTCGCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1781
    |||||
QY 1441 GACCAAGAGGCGCTTATTTGCTGAGCACTACCTCAACCGTGTGATGCTGCTGCGCGTTG 1500
    |||||
DB 1782 GACCAAGAGGCGCTTATTTGCTGAGCACTACCTCAACCGGCGGCTGCTGCTGCTTAAAC 1841
    |||||
QY 1501 CAGGTGTGTGAGGCGCAGTGTACTGTTCCTCACTCCAGGCGCTGTTGTGAGGAGGAGGAGG 1560
    |||||
DB 1842 CAGGTGTGTGAGTTCAGTGTACTGTTCCTCACTCCAGGCGCTGTTGTGAGGAGGAGGAGG 1901
    |||||
QY 1561 CGTTTCGAGGCGCCTTACGATGATGAGGAGTGAATGAGACGAGCGTGTGCTTCTCAAC 1620
    |||||
DB 1902 CGTTTCGAGGCGCCTTACGATGATGAGGAGTGAATGAGACGAGCGTGTGCTTCTTAAAC 1961
    |||||
QY 1621 AACACGCGGCGCGCCACGAGGAGCACTGCTTGTGCTGATGATGATGATGATGATGATG 1680
    |||||
DB 1962 AACACGCGGCGCGCCACGAGGAGCACTGCTTGTGCTGATGATGATGATGATGATGATG 2021
    |||||
QY 1681 ACCAAGACGTTGAGGAGGCGCGCGCGGAGCAATCGGAGGAGGAGGAGGAGGAGGAGGAG 1740
    |||||
DB 2022 ACCAAGACGTTGAGGAGGCGCGCGCGCGGAGCAATCGGAGGAGGAGGAGGAGGAGGAGG 2081
    |||||
QY 1741 TGCCCCACGAGTGTCTTCCGAGGAGCACTCCGAGGAGCACTTACACCAATGCGGTTGCGAG 1800
    |||||
DB 2082 TGCCCCACGAGTGTCTTCCGAGGAGCACTCCGAGGAGCACTTACACCAATGCGGTTGCGAG 2141
    |||||
QY 1801 CCTTGGTTG 1809
    |||||
DB 2142 CCTTGGTTG 2150
    |||||

RESULT 3
ABK91424
ID ABK91424 standard; DNA; 9605 BP.
XX
XX ABK91424;
AC
XX
XX
DT 15-NOV-2002 (first entry)
XX
XX Hepatitis C virus Con 1 isolate DNA mutant 1.
DE
XX
XX HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;
KM hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
KM internal ribosome entry site; IRES; NS5A; HCV replication; mutant.
XX
OS Hepatitis C virus.
OS Synthetic.
XX
XX Key Location/Qualifiers
XX CDS 342..9374
    /tag= a
    /product= "HCV polyprotein"
    /note= "The polyprotein consists of the Core, E1, E2, P7,
    NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
    /tag= b
XX mutation
XX
XX WO200259321-A2.
XX
```

```
PD 01-AUG-2002.
XX
XX 16-JAN-2002; 2002WO-EP000526.
PF
XX
XX 23-JAN-2001; 2001US-0263479P.
PR
XX
XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
PA
XX
XX De Francesco R, Migliaiaccio G, Paonessa G;
PI WPI, 2002-599793/64.
XX
XX
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
PT ribosome entry site (IRES) region, useful in studying HCV replication and
PT expression.
XX
XX Claim 9; Page: 69pp; English.
PS
XX
XX The invention relates to nucleic acid molecules comprising altered HCV
CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
CC internal ribosome entry site (IRES) region coding for one or more NS3;
CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations
CC are detailed in the specification. Also included are (1) an expression
CC vector comprising a nucleotide sequence coding for the altered nucleic
CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a
CC recombinant cell human hepatoma cell comprising the altered nucleic acids
CC ; (3) a recombinant cell produced by introducing into a human hepatoma
CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
CC replicon enhanced cell or which containing a functional HCV replicon; (5)
CC an HCV replicon enhanced cells made in the method; and (6) measuring the
CC ability of a compound to affect HCV activity. The HCV replicons and HCV
CC replicon enhanced cells are useful in studying HCV replication and
CC expression, and HCV and host cell interactions, producing HCV RNA and
CC proteins, and providing a system for measuring the ability of a compound
CC to modulate one or more HCV activities e.g. to discover drugs which may
CC treat HCV mediated diseases such as liver failure, cirrhosis and
CC hepatocellular carcinoma. Note: The present sequence is an HCV replicon Con 1
CC mutant of the invention. Note: The present sequence is not shown in the
CC specification but was created by the index using the HCV sequence
CC appearing as ABK91411 and the information in Claim 9
XX
XX SQ Sequence 9605 BP; 1910 A; 2884 C; 2732 G; 2079 T; 0 U; 0 Other;

Query Match 86.9%; Score 1572.2; DB 6; Length 9605;
Best Local Similarity 91.8%; Pred. No. 0;
Matches 1661; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY 1 ATGAGCAGGAATCTTAAACCTCAAGAAAGAAACCAACGTAACACCAACCGCGCCACAG 60
    |||||
DB 342 ATGAGCAGGAATCTTAAACCTCAAGAAAGAAACCAACGTAACACCAACCGCGCCACAG 401
    |||||
QY 61 GACGTCAAGTTCGCGGCGGTGTCAAGTCTTGTGAGATTACCTGTTCCCGCGCAGG 120
    |||||
DB 402 GACGTCAAGTTCGCGGCGGTGTCAAGTCTTGTGAGATTACCTGTTCCCGCGCAGG 461
    |||||
QY 121 GGGCCGAGGTTGGGAGTGGCGGCGCACTTAGAAGACTTCCGAGCGAGTGCAGCACTGTGGA 180
    |||||
DB 462 GGGCCGAGGTTGGGAGTGGCGGCGCACTTAGAAGACTTCCGAGCGAGTGCAGCACTGTGGA 521
    |||||
QY 181 AGGCGACAACCTATCCCAAGGCTCGCCAGGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
    |||||
DB 522 AGGCGACAACCTATCCCAAGGCTCGCCAGGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 581
    |||||
QY 241 TACCTTGGCGGCTGTATGAGCAAGAGGAGCATGGGAGTGGCGAGAGATGGCTCTCTGACCC 300
    |||||
DB 582 TACCTTGGCGGCTGTATGAGCAAGAGGAGCATGGGAGTGGCGAGAGATGGCTCTCTGACCC 641
    |||||
QY 301 CGTGGCTCCGCGGCTAGTGGGAGGCGCACTGACCCCGGCGGCTAGGTGCGGTAATTTGGGT 360
    |||||
DB 642 CGTGGCTCCGCGGCTAGTGGGAGGCGCACTGACCCCGGCGGCTAGGTGCGGTAATTTGGGT 701
    |||||
QY 361 AAAGTATGATGATACCTTCATGCGGCTTCCGCGAGCTCATGGGAGTACATTCCGCTGTC 420
    |||||
```

Db 702 AAGTATGATACCTTCAGTGGGCTTGCCGATCTATGGGATCATTCGCGCTGTC 761
 QY 421 GGGGCTCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTTCGGAGAC 480
 Db 762 GGGGCCCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTTCGGAGAC 821
 QY 481 GGGGTAACATGACAGGGAAATTACCGGTTGCTTCTATCTCTCTCTGAGT 540
 Db 822 GGGGTAACTATCAACAGGGAAATCTGCCGGTGTCTCTTCTATCTCTCTGAGT 881
 QY 541 TTGCTGCTGTTTGAACATTCAGCTTCGCTTATGAGTGGCAAGTGGGATC 600
 Db 882 TTGCTGCTGTTTGAACATTCAGCTTCGCTTATGAGTGGCAAGTGGGATC 941
 QY 601 TACATGTGACGAGCATGTTCTCAACTCAAGATCTGTATGAGAGAGGAGCATGATC 660
 Db 942 TACATGTGACGAGCATGTTCTCAACTCAAGATCTGTATGAGAGGAGCATGATC 1001
 QY 661 ATGCAACACCCGGGGTGTGCTGTGTGCTGGGAGGATTAATCTCTCCGCTGCTGGT 720
 Db 1002 ATGCAACACCCGGGGTGTGCTGTGTGCTGGGAGGATTAATCTCTCCGCTGCTGGT 1061
 QY 721 GCGCTCACTCCAGCTCGCGGCAAGAGAGCCAGCATCCCACTGGCAATACGAGCG 780
 Db 1062 GCGCTCACTCCAGCTCGCGGCAAGAGAGCCAGCATCCCACTGGCAATACGAGCG 1121
 QY 781 CAGTCGATTTGCTGCTGGGGGCGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 Db 1122 CAGTCGATTTGCTGCTGGGGGCGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1181
 QY 841 TGGGATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 Db 1182 TGGGATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1241
 QY 901 GTACAGAGCTGCAATTTGTTCAATCTATCCCGGCAAGATACAGATGATGATGATGATGAT 960
 Db 1242 GTACAGAGCTGCAATTTGTTCAATCTATCCCGGCAAGATACAGATGATGATGATGATGAT 1301
 QY 961 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
 Db 1302 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1361
 QY 1021 CCACAGCT 1080
 Db 1362 CCACAGCT 1421
 QY 1081 TACTATTTCAATGCTGGGAACTGGGCTAAGGCTTGGGTTGATGATGATGATGATGATGATGAT 1140
 Db 1422 TACTATTTCAATGCTGGGAACTGGGCTAAGGCTTGGGTTGATGATGATGATGATGATGATGAT 1481
 QY 1141 GTTGAAGGGAACTTACAGAGAGGGGAGACAGCGCGCGCGCCAGCGGACTTAA 1200
 Db 1482 GTTGAAGGGAACTTACAGAGAGGGGAGACAGCGCGCGCGCCAGCGGACTTAA 1541
 QY 1201 TCCCTTTTCAACCTGGGCGGCTCAGAAATTCAGCTTGTAAACCAACGCGAGCTG 1260
 Db 1542 TCCCTTTTCAACCTGGGCGGCTCAGAAATTCAGCTTGTAAACCAACGCGAGCTG 1601
 QY 1261 CACATCAAGAGAGCTGCTTGAACCTGCAATGATCTCCCTCAACCTGGGTTCTTGGCGG 1320
 Db 1602 CACATCAAGAGAGCTGCTTGAACCTGCAATGATCTCCCTCAACCTGGGTTCTTGGCGG 1661
 QY 1321 CTGTTCTACAGCAGAGTTCAATGCGTTCGAGTGTCAAGAGGAGCATGGCAGCTGCG 1380
 Db 1662 CTGTTCTACAGCAGAGTTCAATGCGTTCGAGTGTCAAGAGGAGCATGGCAGCTGCG 1721
 QY 1381 CCCATGACAGCTTCAATGAGGGGTGGGTTCCATCACTTAAATGAGTCCAGCGCTTG 1440
 Db 1722 CCCATGACAGCTTCAATGAGGGGTGGGTTCCATCACTTAAATGAGTCCAGCGCTTG 1781
 QY 1441 GACCAAGGCGCTTATGCTGAGCACTCAACCGTGTGATGCGCGGCTTG 1500

Db 1782 GACCAAGGCGCTTATGTTGGCACTACGACCCCGGCGCGGTATCCGATACCCGGGGG 1841
 QY 1501 CAGGTGTGAGCCCACTGATCTGTTCACTCCAGGCCCTGTTGTTGGGAGACGACCGAT 1560
 Db 1842 CAGGTGTGAGTCCAGTGTATGCTTCAACCCCAAGCCCTGCTGTTGGGAGACGACCGAT 1901
 QY 1561 CGTTTGGGCGCCCTGATGATGAGTGGGTTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
 Db 1902 CGTTTGGGCGCCCTGATGATGAGTGGGTTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1961
 QY 1621 AACACGCGCGCGGCAAG 1680
 Db 1962 AACACGCGCGCGGCAAG 2021
 QY 1681 AACACGCGCGCGGCAAG 1740
 Db 2022 AACACGCGCGCGGCAAG 2081
 QY 1741 TGGCCAG 1800
 Db 2082 TGGCCAG 2141
 QY 1801 CCTTGGTTG 1809
 Db 2142 CCTTGGTTG 2150
 RESULT 4
 ABK91429
 ID ABK91429 standard; DNA; 9605 BP.
 AC ABK91429;
 XX
 DT 15-NOV-2002 (first entry)
 XX
 DE Hepatitis C virus Con 1 isolate DNA mutant 6.
 XX
 KM HCV, ds; Con 1, adaptive mutation; liver failure; cirrhosis;
 KM hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
 KM internal ribosome entry site; IRES; NS5A; HCV replication; mutant.
 OS
 OS Hepatitis C virus.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT CDS 342..9374
 FT /*tag= a
 FT /product= "HCV polyprotein"
 FT /note= "The polyprotein consists of the Core, E1, E2, p7,
 FT NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
 FT mutation replace(6859,C)
 FT /*tag= b
 PN MO200259321-A2.
 XX
 PD 01-AUG-2002.
 XX
 PD 16-JAN-2002; 2002WO-EP000526.
 PF
 PR 23-JAN-2001; 2001US-0263479P.
 PA (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
 PA De Francesco R, Migliaccio G, Paonessa G;
 PI WPI; 2002-599793/64.
 DR
 XX
 PT New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
 PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
 PT ribosome entry site (IRES) region, useful in studying HCV replication and
 PT expression.
 PS Claim 9; Page: 69pp; English.

XX The invention relates to nucleic acid molecules comprising altered HCV
CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
CC internal ribosome entry site (IRES) region coding for one or more NS3,
CC NS5a, or EMCV IRES mutations, respectively. The location of the mutations
CC are detailed in the specification. Also included are (1) an expression
CC vector comprising a nucleotide sequence coding for the altered nucleic
CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a
CC recombinant cell human hepatoma cell comprising the altered nucleic acids
CC ; (3) a recombinant cell produced by introducing into a human hepatoma
CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
CC replicon enhanced cell or which containing a functional HCV replicon; (5)
CC an HCV replicon enhanced cells made in the method; and (6) measuring the
CC ability of a compound to affect HCV activity. The HCV replicons and HCV
CC replicon enhanced cells are useful in studying HCV replication and
CC expression, and HCV and host cell interactions, producing HCV RNA and
CC proteins, and providing a system for measuring the ability of a compound
CC to modulate one or more HCV activities e.g. to discover drugs which may
CC treat HCV mediated diseases such as liver failure, cirrhosis and
CC hepatocellular carcinoma. The present sequence is an HCV replicon Con 1
CC mutant of the invention. Note: The present sequence is not shown in the
CC specification but was created by the indexer using the HCV sequence
CC appearing as ABK91411 and the information in Claim 9
XX

SQ Sequence 9605 BP; 1910 A; 2882 C; 2733 G; 2080 T; 0 U; 0 Other;
Query Match 86.9%; Score 1572.2; DB 6; Length 9605;
Best Local Similarity 91.8%; Pred. No. 0;
Matches 1661; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY 1 ATGAGCAGATCTTAACTCAAGAAACCAACGTAACCAACCGCCGACAG 60
DB 342 ATGACACAGATCTTAACTCAAGAAACCAACGTAACCAACCGCCGACAG 401
QY 61 GAGGTCAAGTCCCGGCGGTGTGAGATCGTGTGAGATTACTGTGCGCGAG 120
DB 402 GAGGTCAAGTCCCGGCGGTGTGAGATCGTGTGAGATTACTGTGCGCGAG 461
QY 121 GAGGTCAAGTCCCGGCGGTGTGAGATCGTGTGAGATTACTGTGCGCGAG 180
DB 462 GAGGTCAAGTCCCGGCGGTGTGAGATCGTGTGAGATTACTGTGCGCGAG 521
QY 181 AGGCGCAACCTATCCCAAGGCTGCGACCGCGAGGCGCTGCGCTGCGAG 240
DB 522 AGGCGCAACCTATCCCAAGGCTGCGACCGCGAGGCGCTGCGCTGCGAG 581
QY 241 TACCCCTGGGCGCTTATGAGCAAGAGGCGATGGGTGGCGAGATGGCTCTGCA 300
DB 582 TACCCCTGGGCGCTTATGAGCAAGAGGCGATGGGTGGCGAGATGGCTCTGCA 641
QY 301 CGTGGCTCCCGGCTAGTGGGCGCCCACTGACCCCGGCTGAGGTCGGTAATTGGAGT 360
DB 642 CGTGGCTCCCGGCTAGTGGGCGCCCACTGACCCCGGCTGAGGTCGGTAATTGGAGT 701
QY 361 AAGTCAATGATACCTCATGCGGCTTGGCGACCTTCATGAGGATATTCGCTGCT 420
DB 702 AAGTCAATGATACCTCATGCGGCTTGGCGACCTTCATGAGGATATTCGCTGCT 761
QY 421 GGGGCTCCCTTATGAGGCGCGCTGCGAGGCGCTGCGAGATGGGTGGCGAGGAG 480
DB 762 GGGGCTCCCTTATGAGGCGCGCTGCGAGGCGCTGCGAGATGGGTGGCGAGGAG 821
QY 481 GGGGCTCCCTTATGAGGCGCGCTGCGAGGCGCTGCGAGATGGGTGGCGAGGAG 540
DB 822 GGGGCTCCCTTATGAGGCGCGCTGCGAGGCGCTGCGAGATGGGTGGCGAGGAG 881
QY 541 TTGCTGTCTCTTTTATGACATTCGAGCTTCTTATGAGGAGTGGCGAGATG 600
DB 882 TTGCTGTCTCTTTTATGACATTCGAGCTTCTTATGAGGAGTGGCGAGATG 941
QY 601 TACCATGTCAGAAAGATGCTTCAACGCAAGCATTTGTATGAGGAGCGAGCATGATC 660
DB 942 TACCATGTCAGAAAGATGCTTCAACGCAAGCATTTGTATGAGGAGCGAGCATGATC 1001

QY 661 ATGACACCGCCGCGGTGTGATCCCTGTGTCCGAGAGGATTAATCTCCGCTGCGAGTA 720
DB 1002 ATGACATCCCGGCGGTGTGATCCCTGTGTCCGAGAGGATTAATCTCCGCTGCGAGTA 1061
QY 721 GGGCTCACTCCAGCGCTGCGCGCAAGAGACCGCATCTCCCATCTGCGCAATTCGAGCG 780
DB 1062 GGGCTCACTCCAGCGCTGCGCGCAAGAGACCGCATCTCCCATCTGCGCAATTCGAGCG 1121
QY 781 CAGGTCAATGTCGTGTGGGCGGCGCTTCTGCTCCGCTATGATGAGGAGATCTC 840
DB 1122 CATGTCAATGTCGTGTGGGCGGCGCTTCTGCTCCGCTATGATGAGGAGATCTC 1181
QY 841 TGGCGATGTCGTGTGTGCGGCGGCGCTTCTGCTCCGCTATGATGAGGAGATCTC 900
DB 1182 TGGCGATGTCGTGTGTGCGGCGGCGCTTCTGCTCCGCTATGATGAGGAGATCTC 1241
QY 901 GTACAGAGCTGCAATTTGTAATCTATTCGCGCAAGTATCAGGTCAACGCGATGCTTGG 960
DB 1242 GTACAGAGCTGCAATTTGTAATCTATTCGCGCAAGTATCAGGTCAACGCGATGCTTGG 1301
QY 961 GATATGATGATGAATCTGCTACCTTACAGAGCGGCTAGGTATGCGAGTACTCCGATC 1020
DB 1302 GATATGATGATGAATCTGCTACCTTACAGAGCGGCTAGGTATGCGAGTACTCCGATC 1361
QY 1021 CCACAGAGCTGCTGAGCATGAGTGGCGGCGCCCATCTGAGGAGTCTGCGCGGCTTGGC 1080
DB 1362 CCACAGAGCTGCTGAGCATGAGTGGCGGCGCCCATCTGAGGAGTCTGCGCGGCTTGGC 1421
QY 1081 TACTATTCATGATGAGGAGATCTGAGTCTTGTGTTGATGATCTCTTTTGGCGGC 1140
DB 1422 TACTATTCATGATGAGGAGATCTGAGTCTTGTGTTGATGATCTCTTTTGGCGGC 1481
QY 1141 GTTACCGGAGAACCTTACAGACAGGAGGAGACACACGCGCGCGCCACAGGCGTTACA 1200
DB 1482 GTTACCGGAGAACCTTACAGACAGGAGGAGACACACGCGCGCGCCACAGGCGTTACA 1541
QY 1201 TCCCTTTTACACCTGCGCGCTCAGAAATTCAGGCTTTTAAACACCAACGAGAGCTTG 1260
DB 1542 TCCCTTTTACACCTGCGCGCTCAGAAATTCAGGCTTTTAAACACCAACGAGAGCTTG 1601
QY 1261 CACATCAACAGATCTGCTTGAATCTGCAATGATCTCTTCCAACTGGGCTCTTGGCGCG 1320
DB 1602 CACATCAACAGATCTGCTTGAATCTGCAATGATCTCTTCCAACTGGGCTCTTGGCGCG 1661
QY 1321 CTGTTTCAACAGCAGGTTCAATGCGTCCGAGTCTCAGAGCGGATGCGAGCGCGC 1380
DB 1662 CTGTTTCAACAGCAGGTTCAATGCGTCCGAGTCTCAGAGCGGATGCGAGCGCGC 1721
QY 1381 CCCATTGACAGTTCGATCAGAGGAGTGGGCTCCATCACTTAATTAATGATCCACGCGCTTG 1440
DB 1722 CCCATTGACAGTTCGATCAGAGGAGTGGGCTCCATCACTTAATTAATGATCCACGCGCTTG 1781
QY 1441 GACCAAGGCGCTTATGCTGAGCATAGCAGCTCAACCGTGTGATTCGCGCGCTTG 1500
DB 1782 GACCAAGGCGCTTATGCTGAGCATAGCAGCTCAACCGTGTGATTCGCGCGCTTG 1841
QY 1501 CAGGTGTGCGCAGGTATCTGTTCACTCAAGCGCTGTGTGAGGAGAGCAGCAT 1560
DB 1842 CAGGTGTGCGCAGGTATCTGTTCACTCAAGCGCTGTGTGAGGAGAGCAGCAT 1901
QY 1561 CGTTTGGCGCGCTTACGATCAGATGAGGAGTGAATGAGAGCAGTGTGCTTCTCAAC 1620
DB 1902 CGTTTGGCGCGCTTACGATCAGATGAGGAGTGAATGAGAGCAGTGTGCTTCTCAAC 1961
QY 1621 AACACGCGCGCGCAACGCGGCACTGTTGCGCTGTACATGATGAATAGCAGCGGCTTC 1680
DB 1962 AACACGCGCGCGCAACGCGGCACTGTTGCGCTGTACATGATGAATAGCAGCGGCTTC 2021
QY 1681 ACCAAGAGTGTGAGGAGCGCGCGTGCACATCGGAGGAGTGGCAACAACATTTGATC 1740
DB 2022 ACCAAGAGTGTGAGGAGCGCGCGCGTGCACATCGGAGGAGTGGCAACAACATTTGATC 2081

QY 1741 TGCCGACGAGCTCTCCGAGACATCCCGAGCCACTTACCAATGCGGTGGG 1800
 DB 2082 TGCCGACGAGCTCTCCGAGACATCCCGAGCCACTTACCAATGCGGTGGG 2141
 QY 1801 CCTTGCTT 1809
 DB 2142 CCTTGCTT 2150

RESULT 5
 ABRK91432
 ID ABRK91432 standard; DNA; 9605 BP.
 XX ABRK91432;
 DT 15-NOV-2002 (first entry)
 DE Hepatitis C virus Con 1 isolate DNA mutant 9.
 XX
 XX HCV; de; Con 1; adaptive mutation; liver failure; cirrhosis;
 KM hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
 KW internal ribosome entry site; IRES; NS5A; HCV replication; mutant.
 XX
 OS Hepatitis C virus.
 OS Synthetic.

Key Location/Qualifiers
 FT CDS 342..9374
 FT /tag= a
 FT /product= "HCV polyprotein"
 FT /note= "The polyprotein consists of the Core, E1, E2, p7,
 FT NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
 FT mutation replace(6936,G)
 FT /tag= b

MO200259321.A2.
 PD 01-AUG-2002.
 PF 16-JAN-2002; 2002WO-BE000526.
 XX
 PR 23-JAN-2001; 2001US-0263479P.
 XX
 PA (RICE-) IST-RICERCHIE BIOL MOLECOLARE ANGELETTI.
 XX De Francesco R, Migliaccio G, Paonessa G;
 PI MPI; 2002-599793/64.
 DR
 XX
 PT New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
 PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
 PT ribosome entry site (IRES) region, useful in studying HCV replication and
 PT expression.
 PS
 PS Claim 9; Page; 69pp; English.

The invention relates to nucleic acid molecules comprising altered HCV
 NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
 internal ribosome entry site (IRES) region coding for one or more NS3,
 NS5A, or EMCV IRES mutations, respectively. The location of the mutations
 are detailed in the specification. Also included are (1) an expression
 vector comprising a nucleotide sequence coding for the altered nucleic
 acid, which is transcriptionally coupled to an exogenous promoter; (2) a
 recombinant cell human hepatoma cell comprising the altered nucleic acids
 ; (3) a recombinant cell produced by introducing an HCV (hepatitis C virus)
 cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
 replicon enhanced cell or which containing a functional HCV replicon; (5)
 an HCV replicon enhanced cells made in the method; and (6) measuring the
 ability of a compound to affect HCV activity. The HCV replicons and HCV
 expression, and HCV and host cell interactions, producing HCV RNA and
 proteins, and providing a system for measuring the ability of a compound
 to modulate one or more HCV activities e.g. to discover drugs which may

CC treat HCV mediated diseases such as liver failure, cirrhosis and
 CC hepatocellular carcinoma. The present sequence is an HCV replicon Con 1
 CC mutant of the invention. Note: The present sequence is not shown in the
 CC specification but was created by the indexer using the HCV sequence
 CC appearing as ABRK91411 and the information in Claim 9
 XX
 XX Sequence 9605 BP; 1911 A; 2883 C; 2732 G; 2079 T; 0 U; 0 Other;

Query Match 86.9%; Score 1572.2; DB 6; Length 9605;
 Best Local Similarity 91.8%; Pred. No. 0;
 Matches 1661; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY 1 ATGAGCAGATCTCTAAACCTCAAGAAACCAACGTAACACCAACGCGGCCACAG 60
 DB 342 ATGAGCAGATCTCTAAACCTCAAGAAACCAACGTAACACCAACGCGGCCACAG 401
 QY 61 GAGCTCAAGTCCCGGCGGTGTGATGATGCTGTGATGATGCTGTGCGGCGAG 120
 DB 402 GAGCTCAAGTCCCGGCGGTGTGATGATGATGCTGTGATGATGCTGTGCGGCGAG 461
 QY 121 GGGCCAGAGTGGGTGTGCGCGGCTGAGTGAAGACTTCCGACCGGTGCAACTCGTGA 180
 DB 462 GGGCCAGAGTGGGTGTGCGCGGCTGAGTGAAGACTTCCGACCGGTGCAACTCGTGA 521
 QY 181 AGGCGACACTATCCCAAGGCTGCGGCGGAGGCGGAGGCGGCTGAGCCGCGG 240
 DB 522 AGGCGACACTATCCCAAGGCTGCGGCGGAGGCGGAGGCGGCTGAGCCGCGG 581
 QY 241 TACCTTGGCCCTCTATGAGCAAGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 300
 DB 582 TACCTTGGCCCTCTATGAGCAAGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 641
 QY 301 CGTGGCTCCCGGCTGATGTTGGGGCCCACTGACCCCGGCTGATGTTGGGT 360
 DB 642 CGTGGCTCCCGGCTGATGTTGGGGCCCACTGACCCCGGCTGATGTTGGGT 701
 QY 361 AAGATCATGATACCTTCATGCGGCTGCGGAGCTGATGAGGCTGATGAGGCTGATGAGGCT 420
 DB 702 AAGATCATGATACCTTCATGCGGCTGCGGAGCTGATGAGGCTGATGAGGCTGATGAGGCT 761
 QY 421 GGGCTCCCTTGAAGGGGCGCTGCAAGGCGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 480
 DB 762 GGGCTCCCTTGAAGGGGCGCTGCAAGGCGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 821
 QY 481 GGGCTGAATGATGCAAGGAAATTTACCGGCTGCTTCTGATCTCTCTGCTGCTGCT 540
 DB 822 GGGCTGAATGATGCAAGGAAATTTACCGGCTGCTTCTGATCTCTCTGCTGCTGCT 881
 QY 541 TTGCTGCTGCTTGAAGCAATTCAGCTTCCGCTTGAAGTGGAGTGGAGTGGAGTGGAGTGG 600
 DB 882 TTGCTGCTGCTTGAAGCAATTCAGCTTCCGCTTGAAGTGGAGTGGAGTGGAGTGGAGTGG 941
 QY 601 TACCATGTCAGCAAGATGCTCAATCTCAAGCATGCTGTAAGAGCAAGCGGACATGATC 660
 DB 942 TACCATGTCAGCAAGATGCTCAATCTCAAGCATGCTGTAAGAGCAAGCGGACATGATC 1001
 QY 661 ATGACACCCCGGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 DB 1002 ATGACACCCCGGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1061
 QY 721 GGGCTCACTCCAGCGTGGCGGCAAGAGCGGAGCATCCCACTGCGCAATTAAGACGCG 780
 DB 1062 GGGCTCACTCCAGCGTGGCGGCAAGAGCGGAGCATCCCACTGCGCAATTAAGACGCG 1121
 QY 781 CAGCTGATTTGCTGTTGGGCGGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 DB 1122 CAGCTGATTTGCTGTTGGGCGGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1181
 QY 841 TGGGATGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 DB 1182 TGGGATGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1241
 QY 901 GTCAGAGCTGCAATGTTCAATCTATCCGCGCACGTAACGATCACCGCATGCTGCTGCTGCT 960

```
Db      1242  GATACAGAGCTGCAATTTGCTCAATATATCCCGGCGACAGGTACACCGTATGCTTGG 1301
OY      961  GATATGATGATGAACTGTCACCTACAGAGCCCTAGTGGTATCGAGCTACTCCGGATC 1020
Db      1302  GATATGATGATGAACTGTCACCTACAGAGCCCTAGTGGTATCGAGCTACTCCGGATC 1361
OY      1021  CCACAAGCTGTGCTGAGACATGCTGCGGGGGCCCACTGCGGAGTCTTGGCGGCTTGGC 1080
Db      1362  CCACAAGCTGTGCTGAGATGCTGCGGGGGCCCACTGCGGAGTCTTGGCGGCTTGGC 1421
OY      1081  TACTATTCATGCTGAGGAACTGGGCTAAGTCTTGGTGTGATGCTTACTTTTGGCCGC 1140
Db      1422  TACTATTCATGCTGAGGAACTGGGCTAAGTCTTGGTGTGATGCTTACTTTTGGCCGC 1481
OY      1141  GTTGAACGGGGAACTTACACAGACAGGGGGGACACACGCGCGCGCCACGCGGCTTACA 1200
Db      1482  GTTGAACGGGGAACTTATGTGACAGGGGGGACGATGGCCAAAAACACCTCGGATTTAG 1541
OY      1201  TCCCTCTTACACCTGCGGCGGCTCAGAAATCCAGCTTTGTAACACCAACGCGAGCTGG 1260
Db      1542  TCCCTCTTACACCGGGGTCATCCCGAATAATCCAGCTTTGTAACACCAACGCGAGCTGG 1601
OY      1261  CACATCAACAGAACTGCTTGAATCTGCAATGACTCCCTCAAACTGGGTTCTTGGCCGG 1320
Db      1602  CACATCAACAGAACTGCTTGAATCTGCAATGACTCCCTCAAACTGGGTTCTTGGCCGG 1661
OY      1321  CTGTTCTACACGACAGGTTTCAATGCGTCCGAGTGTCTGAGCGGACATGGCGAGCTGGC 1380
Db      1662  CTGTTCTACAGTGCACAGTTTCAATGCTGATGATCCCGAGCGCATGGCCAGCTGGCAC 1721
OY      1381  CCCATTGACAGATTGCTGATCAGGGGTGGGGTCCCATCTATATATGATGCCACGGCTTG 1440
Db      1722  CCCATTGACAGCTTGGCTGCTGAGGGGTGGGGGCCATCTTACATATGATGACACAGCTGG 1781
OY      1441  GACCAAGGCGCCCTATTTGCTGAGACCTACGCACTGTGTGATGCTGCGCCGCTTG 1500
Db      1782  GACCAAGGCGCCCTATTTGCTGAGACCTACGCACTGTGTGATGCTGCGCCGCTTG 1841
OY      1501  CAGTGTGTGCGCCCACTGACTGTTTCACTCCAGAGCCCTGTTGGTGGGAGACGACCGAT 1560
Db      1842  CAGTGTGTGCGCCCACTGACTGTTTCACTCCAGAGCCCTGTTGGTGGGAGACGACCGAT 1901
OY      1561  CGTTTGGGCGCCCTTACGTAAGATGGGGTGAATGAGACGAGCTGTGCTTCTCAAC 1620
Db      1902  CGTTTGGGCGCCCTTACGTAAGATGGGGTGAATGAGACGAGCTGTGCTTCTCAAC 1961
OY      1621  AACAGCGCGCGCGCCACGCGGCGCACTGTTCCGCTGTATCATGATGATAGCACTGGCTTC 1680
Db      1962  AACAGCGCGCGCGCCACGCGGCGCACTGTTCCGCTGTATCATGATGATAGCACTGGCTTC 2021
OY      1681  ACCAAGAGCTGTGGGGGCGCCCGCTGCAACATCGGGGGGCTGGGCAACCACTTTGATC 1740
Db      2022  ACCAAGAGCTGTGGGGGCGCCCGCTGTAACATCGGGGGGATGGGCAATTAACCTTTGACC 2081
OY      1741  TGCCCCAGGAGCTGTTCCGGAAGCATCCGAGGCGCACTTACCACAATGCGGTTGGGG 1800
Db      2082  TGCCCCAGGAGCTGTTCCGGAAGCATCCGAGGCGCACTTACCACAAGTGTGTTGGGG 2141
OY      1801  CTTGGTTG 1809
Db      2142  CTTGGTTG 2150
```

RESULT 6
ABK91411
ID ABK91411 standard; DNA; 9605 BP.
XX
XX ABK91411;
AC
XX
DT 15-NOV-2002 (first entry)
XX
XX Hepatitis C virus Con 1 isolate DNA.

```
XX      HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;  
KW      hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;  
KW      internal ribosome entry site; IRES; NS5a; HCV replication.  
OS      Hepatitis C virus.  
XX  
FH      Key  
FT      CDS  
FT      Location/Qualifiers  
FT      /tag=a  
FT      /product="HCV polypeptide"  
FT      /note="The polypeptide consists of the Core, E1, E2, p7,  
FT      NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"  
PN      WO200259321-A2.  
XX  
PD      01-AUG-2002.  
XX  
XX      16-JAN-2002; 2002WO-EP000526.  
XX  
XX      23-JAN-2001; 2001US-0263479P.  
XX  
XX      (RICE-) IST RICEBCH BIOLOGIC MOLECULAR ANGLETTI.  
PI      De Francesco R, Migliaccio G, Paonessa G;  
XX  
XX      WPI; 2002-599793/64.  
DR      P-PSDB; ABG32451.  
XX  
XX      New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV  
PT      NS5 encoding region, or encephalomyocarditis virus (EMCV) internal  
PT      ribosome entry site (IRES) region, useful in studying HCV replication and  
PT      expression.  
XX  
XX      Claim 9; Page 36-39; 69pp; English.  
PS  
XX      The invention relates to nucleic acid molecules comprising altered HCV  
XX      NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)  
XX      internal ribosome entry site (IRES) region coding for one or more NS3,  
XX      NS5a, or EMCV IRES mutations, respectively. The location of the mutations  
XX      are detailed in the specification. Also included are (1) an expression  
XX      vector comprising a nucleotide sequence coding for the altered nucleic  
XX      acids, which is transcriptionally coupled to an exogenous promoter; (2) a  
XX      recombinant cell human hepatoma cell comprising the altered nucleic acids  
XX      ; (3) a recombinant cell produced by introducing into a human hepatoma  
XX      cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)  
XX      replicon enhanced cell or which containing a functional HCV replicon; (5)  
XX      an HCV replicon enhanced cells made in the method; and (6) measuring the  
XX      ability of a compound to affect HCV activity. The HCV replicons and HCV  
XX      replication enhanced cells are useful in studying HCV replication and  
XX      expression, and HCV and host cell interactions, producing HCV RNA and  
XX      proteins, and providing a system for measuring the ability of a compound  
XX      to modulate one or more HCV activities e.g. to discover drugs which may  
XX      treat HCV mediated diseases such as liver failure, cirrhosis and  
XX      hepatocellular carcinoma. The present sequence is the HCV replicon Con 1,  
XX      used as a basis for the adaptive mutations of the invention  
SQ      Sequence 9605 BP; 1910 A; 2883 C; 2733 G; 2079 T; 0 U; 0 Other;  
OY      Query Match 86.9%; Score 1572.2; DB 6; Length 9605;  
OY      Best Local Similarity 91.8%; Pred. No. 0;  
Db      Matches 1661; Conservative 0; Mismatches 148; Indels 0; Gaps 0;  
OY      1 ATGAGCAGGATCCCTAAACCTCAAGAAACCAAGCTACACCAACCGCGGCCACAG 60  
Db      342 ATGAGCAGGATCCCTAAACCTCAAGAAACCAAGCTACACCAACCGCGGCCACAG 401  
OY      61 GACGTCAAGTCCCGGGCGGTGTGATCGTGTGAGTGTGATCTTGTCCCGCGCAG 120  
Db      402 GACGTCAAGTCCCGGGCGGTGTGATCGTGTGAGTGTGATCTTGTCCCGCGCAG 461  
OY      121 GCGCCAGGTTGGTGTGCGCGCACTAGAAAGCTTCCGAGCGGTGCAACTCTGTGGA 180
```

Db	442	GGCCCCAGAGTTGGGTTGGTGGCGGCACTAGAAAGACTTCCGAGCGGTGCGCAACTCTGTGGA	521
Oy	181	AGCGGCAACCTTATCCCAAGGCTTCGACGCCCCGAAGGGCAGGCTGGGCTCAGCCGGG	240
Db	522	AGGGACAACCTATCCCAAGGCTCGACAGCCCGAGGTAAAGGCTTGGGCTCAGCCGGG	581
Oy	241	TACCTTGGGCCCCCTCTATAGGCAAGAGGGCAATGGGGTGGGAGAGAATGCTCTCTCACC	300
Db	582	TACCCCTGGCCCCCTCTATAGGCAAGAGGGCTTGGGGTGGGAGAGATGCTCTCTGCAACC	641
Oy	301	CGTGGCTCCGGGCTAGTTGGGGCCCCCACTAACCCCCGAGCTTAAGTGCAGTAATTTGGGT	360
Db	642	CGTGGCTCCGGCTAGTTGGGGCCCCCAAGAACCCCCGAGCTTAAGTGCAGTAATTTGGGT	701
Oy	361	AAAGTCATGATACCTTCACATGCGGCTTTCGCGACTCATGGGGTACATTTCGCTCGTC	420
Db	702	AAGGTCATGATACCTTCACGTCGGGCTTCGCGCGATCTCATGGGGTACATTTCGCTCGTC	761
Oy	421	GGCGCTCCCTTAAAGGGGGCGCTGCGCAGGGGCCCTGGCGCATGCGCTCCGGGTTCTGAAGAC	480
Db	762	GGCGCCCCCTTAAAGGGGGCGCTGCGCAGGGGCCCTGGCGCATGCGCTCCGGGTTCTGAAGAC	821
Oy	481	GGCGGAACTATGCAACAGGGGAATTTACCGGGTTCCTTCTCATTTCTCTGGCT	540
Db	822	GGCGGAACTATGCAACAGGGGAATTTGCGCGGTTCTCTTTCATTTCTTTTGACT	881
Oy	541	TTGCTGCTCTGTTTGACCATTCACGCTTCGCGCTTATGAAGTGGCMACTGTCGGGATC	600
Db	882	TTGCTGCTCTGTTTGACCATTCAGACTTCGCGCTTATGAAGTGGCMACTGTCGGGATG	941
Oy	601	TAGCATGTCAAGAACGATTTGCTCCAACTCAAGCATCGTGAAGAGACAGCGGACATGATC	660
Db	942	TACCATGTCAAGAACGACTGCTCCAACCCAGCATTTGTATGAGGCGAGCGAATGATC	1001
Oy	661	ATGCAACACCCCGGGTGTGTGCGCTGTGCTCCGGGAGGGTAATTCCTCCGCTGCGGGTA	720
Db	1002	ATGCAATACCCCGGGTGTGTGCGCTGTGCTGCGGGAACAACCTCTCCGCTGCTGGGTA	1061
Oy	721	GCCTCATCTCCACGCTGCGCGGCAAGAGCGCAGCATCCCACTGCGACAATAAGACGC	780
Db	1062	GCGCTCATCTCCACGCTGCGCGGCAAGAGCGTATGCGTCCCACTGACGACATAGACGC	1121
Oy	781	CACGTGATTTGCTGTTGGGGCGGCTGCTTGTGCTCCGCTATGTACGTGGGGATTC	840
Db	1122	CATGTGATTTGCTGTTGGGGCGGCTGCTCTGTGCTCCGCTATGTACGTGGGGATTC	1181
Oy	841	TGCGGATTCGTTTCTCTGTCCTCAGCTGTACACTTCGCGCGCGGCAATCAAGCG	900
Db	1182	TGCGGATTCGTTTCTCTGTCGCCAGCTGTTCACCTTCTGCGCTCGCGGCAAGAGCA	1241
Oy	901	GTACAGACTCAATTGTTCAATCATCCCGGCAAGTACAGGTCAACGATGGCTTGG	960
Db	1242	GTACAGACTCAATTGCTCAATATATCCCGGCAAGTACAGGTCAACCGTATGGCTTGG	1301
Oy	961	GATATGATGATGAACGTGTCACCTTACAGACGCCCTTAGTGTATGCGAGCTACTCGGATC	1020
Db	1302	GATATGATGATGAACGTGTCACCTTACAGACGCCCTTAGTGTATGCGAGTACTCGGATC	1361
Oy	1021	CCACAGCTGTGCTGAGACATAGTGGCGGGGGCCCACTGGGGAGTCTTGGCGGGCTTGGC	1080
Db	1362	CCACAGCTGTGCTGAGATAGTGGTGGCGGGGGCCCACTGGGGAGTCTTGGCGGGCTTGGC	1421
Oy	1081	TACTATTCATGATGGGAGACGTGGGCTTAAGGCTGTTGTTGATGCTACTCTTTGGCGGC	1140
Db	1422	TACTATTCATGATGGGAGACGTGGGCTTAAGGCTTCTGATTTGTATGCTACTCTTTGGCGGC	1481
Oy	1141	GTTGACGGGGAACTTACACGACAGAGGGGGACACACGGCGCGCGCCACAGGGCTTACA	1200
Db	1482	GTTGACGGGGAACTTATGTGACAGGGGGAGCATGGCCAAAAAACACCTCGGGATTACG	1541
Oy	1201	TCCCTCTTCAACACTCGGGCGGCTCAGAAAATCCAGCTTGTAAACACMAAGGACGCTGG	1260
Db	1542	TCCCTCTTTCACCCGGGTCATCCCAAGAAATCCAGCTTGTAAACACMAAGGACGCTGG	1601

Oy		1261	CACATCAACAAGCATGCGCTTAACTGGCAATGACTCCCTCCAAACTGGGTTCCGTTCGCGC	1320
Oy		1262	CACATCAACAAGCATGCGCTTAACTGGCAATGACTCCCTCCAAACTGGGTTCCGTTCGCGC	1361
Dd		1602	CACATCAACAAGCATGCGCTTAACTGGCAATGACTCCCTCCAAACTGGGTTCCGTTCGCGC	1661
Oy		1321	CTGTTCTAACAAGCATGCGCTTAACTGGCAATGACTCCCTCCAAACTGGGTTCCGTTCGCGC	1380
Dd		1662	CTGTTCTAACAAGCATGCGCTTAACTGGCAATGACTCCCTCCAAACTGGGTTCCGTTCGCGC	1721
Oy		1381	CCCATTAACAGTTTCGATCAGGGGTGGGGTCCCATTACTTATAATGATGCCAAGGCTTG	1440
Dd		1722	CCCATTAACAGTTTCGATCAGGGGTGGGGTCCCATTACTTATAATGATGCCAAGGCTTG	1781
Oy		1441	GACCAGAAGCCCTTAATGCTGGCACTACGACCTTCMAACGATGTGGTATCTGCCCCGCTTG	1500
Dd		1782	GACCAGAAGCCCTTAATGCTGGCACTACGACCTTCMAACGATGTGGTATCTGCCCCGCTTG	1841
Oy		1501	CAGGTGTGTGGCCCAAGTGAATGTTTCACTTCMAAGCCCTGTTGTGTGGGAGCACCGAT	1560
Dd		1842	CAGGTGTGTGGCCCAAGTGAATGTTTCACTTCMAAGCCCTGTTGTGTGGGAGCACCGAT	1901
Oy		1561	CGTTTCGGCGCCCTTAAGTACATATGGGGTGAATATGACAGCACGTGCTCTTCTAAC	1620
Dd		1902	CGTTTCGGCGCCCTTAAGTACATATGGGGTGAATATGACAGCACGTGCTCTTCTAAC	1961
Oy		1621	AACACGGCGGCCCAAGCGGCGCACTGGTTCGGCTGATATGATGATTAAGACACCGGTTG	1680
Dd		1962	AACACGGCGGCCCAAGCGGCGCACTGGTTCGGCTGATATGATGATTAAGACACCGGTTG	2021
Oy		1681	ACCAAGACGTGTGGGGGGCCCCCGTGCACATCGGGGGGGGTTCGGCAACACCTTTGATC	1740
Dd		2022	ACCAAGACGTGTGGGGGGCCCCCGTGCACATCGGGGGGGGTTCGGCAACACCTTTGATC	2081
Oy		1741	TGCCCCACGACCTGCTTCCGGAACATCCGAGGCCACTTAACCAATCGGTTCCGGG	1800
Dd		2082	TGCCCCACGACCTGCTTCCGGAACATCCGAGGCCACTTAACCAATCGGTTCCGGG	2141
Oy		1801	CCTTGATTG 1809	
Dd		2142	CCTTGATTG 2150	
RESULT 7				
ABK91430				
ID	ABK91430 standard; DNA; 9605 BP.			
XX				
AC	ABK91430;			
XX				
DT	15-NOV-2002 (first entry)			
XX				
DE	Hepatitis C virus Con 1 isolate DNA mutant 7.			
XX				
KM	HCV, db; Con 1; adaptive mutation; liver failure; cirrhosis;			
KW	hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;			
KM	internal ribosome entry site; IRES; NS5A; HCV replication; mutant.			
XX				
OS	Hepatitis C virus.			
XX				
Key				
FH	Location/Qualifiers			
FT	CDS 342..9374			
FT	/*tag= a			
FT	/product= "HCV polypotein"			
FT	/note= "The polypotein consists of the Core, E1, E2, P7,			
FT	NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"			
FT	mutation replace(6931,C)			
XX				
PN	WO200259321-A2.			
XX				
PD	01-AUG-2002.			
XX				
PF	16-JAN-2002; 2002WO-EP000526.			

XX 23-JAN-2001; 2001US-0263479P.
PR (RICE-) IST RIGERCHE BIOL MOLECOLARE ANGELETTI.
PA De Francesco R, Migliaccio G, Paonessa G;
XX WPI; 2002-599793/64.
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
PT ribosome entry site (IRES) region, useful in studying HCV replication and
PT expression.
XX Claim 9; Page; 69pp; English.

XX The invention relates to nucleic acid molecules comprising altered HCV
CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
CC internal ribosome entry site (IRES) region coding for one or more NS3,
CC NS5a, or EMCV IRES mutations, respectively. The location of the mutations
CC are detailed in the specification. Also included are (1) an expression
CC vector comprising a nucleotide sequence coding for the altered nucleic
CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a
CC recombinant cell human hepatoma cell comprising the altered nucleic acids
CC; (3) a recombinant cell produced by introducing into a human hepatoma
CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
CC replicon enhanced cell or which containing a functional HCV replicon; (5)
CC an HCV replicon enhanced cells made in the method; and (6) measuring the
CC ability of a compound to affect HCV activity. The HCV replicons and HCV
CC replicon enhanced cells are useful in studying HCV replication and
CC expression, and HCV and host cell interactions, producing HCV RNA and
CC proteins, and providing a system for measuring the ability of a compound
CC to modulate one or more HCV activities e.g. to discover drugs which may
CC treat HCV mediated diseases such as liver failure, cirrhosis and
CC hepatocellular carcinoma. The present sequence is an HCV replicon Con 1
CC mutant of the invention. Note: The present sequence is not shown in the
CC specification but was created by the indexer using the HCV sequence
CC appearing as ABK91411 and the information in Claim 9
XX SQ

Sequence 9605 BP; 1910 A; 2882 C; 2733 G; 2080 T; 0 U; 0 Other;

Query Match 86.9%; Score 1572.2; DB 6; Length 9605;

Best Local Similarity 91.8%; Pred. No. 0;

Matches 1661; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY	1	ATGAGCAGCAATCTCAAAACCTCAAGAAACCAACGTAACCAACCGCGCCACAG	60
DB	342	ATGAGCAGCAATCTCAAAACCTCAAGAAACCAACGTAACCAACCGCGCCACAG	401
QY	61	GACGTCAAGTCCCGGCGGTGTGATCGTGTGTGAGTTTACCTGTGTGCGCGCAG	120
DB	402	GACGTCAAGTCCCGGCGGTGTGATCGTGTGTGAGTTTACCTGTGTGCGCGCAG	461
QY	121	GGCCCCAGGTGTGTGTGCGCGGACTAGAAAGACTTCGACCGGTGCAACCTGTGGA	180
DB	462	GGCCCCAGGTGTGTGTGCGCGGACTAGAAAGACTTCGACCGGTGCAACCTGTGGA	521
QY	181	AGGCGCAACCTATCCCAAGGCTGCGCAGCCGAGGCGGAGGCGCTGAGCCCGG	240
DB	522	AGGCGCAACCTATCCCAAGGCTGCGCAGCCGAGGCGGAGGCGCTGAGCCCGG	581
QY	241	TACCCCTTGCGCTCTTATGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	300
DB	582	TACCCCTTGCGCTCTTATGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	641
QY	301	CGTGTGCTCCCGGCTGTGTGTGCGCGGCTGAGCCCGGCTGAGCTGTATTTGGGT	360
DB	642	CGTGTGCTCCCGGCTGTGTGTGCGCGGCTGAGCCCGGCTGAGCTGTATTTGGGT	701
QY	361	AAAGTCATGATACCTTCATGATGCGGCTTGCAGCACTGATGAGGAGTATTCGCTGTC	420
DB	702	AAAGTCATGATACCTTCATGATGCGGCTTGCAGCACTGATGAGGAGTATTCGCTGTC	761

QY	421	GGCGCTCCCTTAGAGGGGCGTCCAGAGGCGCTGTGCGCATGCGGTCCGGGTTCTGAGAGAC	480
DB	762	GGCGCTCCCTTAGAGGGGCGTCCAGAGGCGCTGTGCGCATGCGGTCCGGGTTCTGAGAGAC	821
QY	481	GGCGTGAATATGCAACAGGGAATTTACCGGCTGTCTTCTATTTCTCTTGGCT	540
DB	822	GGCGTGAATATGCAACAGGGAATTTACCGGCTGTCTTCTATTTCTCTTGGCT	881
QY	541	TTGCTGCTCTGTTTACCATTCAGCTTCGCGCTTATGAGTGGCGCAAGTGTCCGGAGTC	600
DB	882	TTGCTGCTCTGTTTACCATTCAGCTTCGCGCTTATGAGTGGCGCAAGTGTCCGGAGTC	941
QY	601	TACCATGTACGAACGATGTGCTCAACTCAAGCATGTGTACGAGACAGCGGACATGATC	660
DB	942	TACCATGTACGAACGATGTGCTCAACTCAAGCATGTGTACGAGACAGCGGACATGATC	1001
QY	661	ATGCAACACCCCGGGGTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG	720
DB	1002	ATGCAATACCCCGGGGTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG	1061
QY	721	GGGCTCACTCCGACGCTGCGGCGGCAAGGAGCGGAGCATCCGCACTGCGCAATACGACGC	780
DB	1062	GGGCTCACTCCGACGCTGCGGCGGCAAGGAGCGGAGCATCCGCACTGCGCAATACGACGC	1121
QY	781	CACGTGATTTGCTGTGTGCGGCGGCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG	840
DB	1122	CATGTCATTTGCTGTGTGCGGCGGCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG	1181
QY	841	TGCGGATGCTGTTTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG	900
DB	1182	TGCGGATGCTGTTTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG	1241
QY	901	GTAAGAGCTCAATTTGCTCAATCTATCCCGGCAAGTATGAGTCAACCGATGAGCTTGG	960
DB	1242	GTAAGAGCTCAATTTGCTCAATCTATCCCGGCAAGTATGAGTCAACCGATGAGCTTGG	1301
QY	961	GATATGATGATGAATCTGTGCTCACTACAGACGCTGTGTATGCGACGTAATCTCGATC	1020
DB	1302	GATATGATGATGAATCTGTGCTCACTACAGACGCTGTGTATGCGACGTAATCTCGATC	1361
QY	1021	CCACAAGCTGTGTGTGATGATGTGTGCGGCGGCGGCACTGTGCGGCGGCTTGGC	1080
DB	1362	CCACAAGCTGTGTGTGATGATGTGTGCGGCGGCGGCACTGTGCGGCGGCTTGGC	1421
QY	1081	TACTATTCATGATGAGGGAATGAGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1140
DB	1422	TACTATTCATGATGAGGGAATGAGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1481
QY	1141	GTTGACGGGGAACCTTACACGACAGGAGGAGACACAGCGCGCGCGCCACGCGCTTACA	1200
DB	1482	GTTGACGGGGAACCTTATGTGACAGGAGGAGACAGATGCGCAAAAACACCTTCGGGATTACG	1541
QY	1201	TCCCTTTCAACCTGTGCGGCGGCTCAAGAAATCCAGCTTGTAAACCAAGGAGGAGCTGG	1260
DB	1542	TCCCTTTTCAACCTGTGCGGCGGCTCAAGAAATCCAGCTTGTAAACCAAGGAGGAGCTGG	1601
QY	1261	CACATCAACAGAACTGCTTGAATGATGATGCTTCCCAATGAGGTTCTTTCGCGCG	1320
DB	1602	CACATCAACAGAACTGCTTGAATGATGATGCTTCCCAATGAGGTTCTTTCGCGCG	1661
QY	1321	CTGTTCATACGACAGGTTCAATGCTGTGAGTGTCAAGGCGATGAGCGAGTGTGCGC	1380
DB	1662	CTGTTCATACGACAGGTTCAATGCTGTGAGTGTCAAGGCGATGAGCGAGTGTGCGC	1721
QY	1381	CCCATTTACAGTTTCAATGAGGAGGAGTGTGCTTCACTTATATGAGTCCCAAGGCTTGG	1440
DB	1722	CCCATTTACAGTTTCAATGAGGAGGAGTGTGCTTCACTTATATGAGTCCCAAGGCTTGG	1781
QY	1441	GACCAAGGAGCTTATGCTGTGCACTACGACCTCAACCGTGTGTGTGTGTGTGTGTGTGT	1500
DB	1782	GACCAAGGAGCTTATGCTGTGCACTACGACCTCAACCGTGTGTGTGTGTGTGTGTGTGT	1841
QY	1501	CAGGTGTGTGCGCAGGTGATCTGTTTCACTCAAGCCCTGTGTGTGTGTGTGTGTGTGT	1560

Db	1002	ATGCATACCCCGGGTGGGTGCTGCCTCGCTTGGGAGAACTCTCTCCCGCTGCTGGGTA	1061
QY	721	GGCGTCACTCCCAACGCTCGCGGCAAGAGCGCAGACATCCCACTGCGACAAATGACGC	780
Db	1062	GGCGTCACTCCCAACGCTCGCGGCAAGAGCGTACGCTCCCACTGACGATGACGC	1121
QY	781	CACGTGCAATTTGCTCGTTGGGGCGGCTTCTTCTCTCCGTTATGTAGTGGGGATCTC	840
Db	1122	CATGTGCAATTTGCTCGTTGGGGCGGCTCTCTCTGCTCCGTTATGTAGTGGGGATCTC	1181
QY	841	TGCGGATCTGTATTTCTCGTCACTCAGCTGTTCACCTTCTCGCGCTCGCGACATCAGACG	900
Db	1182	TGCGGATCTGTATTTCTCGTCACTCAGCTGTTCACCTTCTCGCGCTCGCGACAGACAC	1241
QY	901	GTAACAGACTGCATATTGTTCAATCTATCCGCGCAAGTATCAGGTCAACGAGTGGCTTGG	960
Db	1242	GTAACAGACTGCATATTGTTCAATCTATCCGCGCAAGTATCAGTATGGCTTGG	1301
QY	961	GATATGATGATGAATCGTCACTACAGACCTTATGTGTATGTGCGAGCTACTCCGATC	1020
Db	1302	GATATGATGATGAATCGTCACTACAGACCTTATGTGTATGTGCGAGCTACTCCGATC	1361
QY	1021	CCACAAGCTGTGCTGGGACATGGTGGCGGGGCGCACTGGGGAGTCTGGCGGGCTTGGC	1080
Db	1362	CCACAAGCTGTGCTGGGACATGGTGGCGGGGCGCACTGGGGAGTCTTACGGCGGGCTTGGC	1421
QY	1081	TACTATTCATGTGTGGGACCTGGGCTTAAGTCTTGTGTATGTACTCTTTTGGCGGC	1140
Db	1422	TACTATTCATGTGTGGGACCTGGGCTTAAGTCTTGTGTATGTACTCTTTTGGCGGC	1481
QY	1141	GTTGACGGGGAACTTTACACGACAGGGGGGACACAGCGCGCGCGCTTACAGGGCTTACA	1200
Db	1482	GTTGACGGGGAACTTTAGTACAGGGGGGACGATGGCAAAAACACCTTCGGGATTAACG	1541
QY	1201	TCCGCTTTACACCTGGGCGCGGCTCAGAAAAATCCAGCTTGTAAAACCAACGGGAGCTGG	1260
Db	1542	TCCGCTTTTACACCGGGCTCATCCAGAAAAATCCAGCTTGTAAAACCAACGGGAGCTGG	1601
QY	1261	CACATCAACAGAACTGCTTGAATGTCAATGACTCCCTCCAAACTGGGTTCTTTGCCGCG	1320
Db	1602	CACATCAACAGAACTGCCCTGACCTGCAATGTACTCCCTCAACCTGGGTTCTTTGCCGCG	1661
QY	1321	CTGTCTTACACGACACAGGTTCAATGCGTCCGATGCTGACAGCGCATGGCAGCTGCCGC	1380
Db	1662	CTGTCTTACGTCACAGAGTTCAACTCATCTGGAATGCCACAGCGCATGGCAGCTGCAGC	1721
QY	1381	CCCATATGACAGTTGCGATCAGGGGGGGGGTCCCATATTTATATAGTCCCAAGGCTTG	1440
Db	1722	CCCATATGACAGGTTGCTCATAGGGGGGGGGCCCATCACTTCAATATGATCAACAGCTCG	1781
QY	1441	GACCAAGAGGCTTATGTCTGGCACTACCACTCAACGATGCTGTGATCGTCCCGCTTG	1500
Db	1782	GACCAAGAGGCTTATGTGTGGCACTACGCAACCCGCGCGCTGCGGATGTGATCCGCGGCG	1841
QY	1501	CAGGTGTGTGGCCAGTGTACTGTTTCACTTCAAGCCCTGTTGTGTGGGACGACCGAT	1560
Db	1842	CAGGTGTGTGGCCAGTGTACTGCTTCAACCCCAAGCCTGTGCTGTGTGGGACGACCGAC	1901
QY	1561	GGTTTGGGCGCCCTTACGTAAGAGTGGGGTGAATGAGACGACGCTGCTTCTCAAC	1620
Db	1902	GGTTTGGGCGCTTCTTACGTAAGAGTGGGGGAGATGAGACGACGCTGCTTCTTAAAC	1961
QY	1621	AACACGCGGCGCGCAACGGGGCAACTGGTTCGGCTGTACATGGAATGAAAGCAGGGGTTG	1680
Db	1962	AACACGCGGCGCGCGCAAGGCAACTGGTTCGGCTGTACATGGAATGAAAGCAGGGGTTG	2021
QY	1681	ACCAAGACGTGTGGGGGCGCCCTCGTGCAACTGGGGGGGATCGGCAACAACTTTGATC	1740
Db	2022	ACCAAGACGTGTGGGGGCGCCCTCGTGTAACATCGGGGGGATCGGCAATTAACCTTGACC	2081
QY	1741	TGCCCCACGACCTGTTCCCGAAGATCCCGAGGCACTTACACCAATGCGGTTTGGGG	1800
Db	2082	TGCCCCACGACCTGTCTCCGAAGACCCCGAGGGCACTTACACCAATGTGTGTGGGG	2141

Qy	1801	CCTTGCTG	1809
Db	2142	CCTTGCTG	2150
RESULT 9			
ABK91425			
ID	ABK91425	standard; DNA; 9605 BP.	
AC	ABK91425;		
XX			
XX			
DT	15-NOV-2002	(first entry)	
XX			
DE	Hepatitis C virus Con 1 isolate DNA mutant 2.		
XX			
KW	HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;		
KM	hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;		
KW	internal ribosome entry site; IRES; NS5A; HCV replication; mutant.		
XX			
OS	Hepatitis C virus.		
XX			
OS	Synthetic.		
XX			
Key	Location/Qualifiers		
XX	CDS	342..9374	
FT		/*tag= a	
FT		/product= "HCV polypeptide"	
FT		/note= "The polypeptide consists of the Core, E1, E2, P1,	
FT		NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"	
FT		replace(3946,A)	
FT	mutation	/*tag= b	
XX			
PM	WO200259321-A2.		
XX			
PD	01-AUG-2002.		
XX			
XX			
PF	16-JAN-2002; 2002WO-EP000526.		
XX			
PR	23-JAN-2001; 2001US-0263479P.		
XX			
PA	(RICE-) IST RICE RCH E BIOL MOLECULAR E ANGELETTI.		
XX			
PI	De Francesco R, Migliaccio G, Paonessa G;		
XX			
DR	WPI; 2002-5599793/64.		
XX			
XX			
PT	New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV		
PT	NS5 encoding region, or encephalomyocarditis virus (EMCV) internal		
PT	ribosome entry site (IRES) region, useful in studying HCV replication and		
PT	expression.		
XX			
PS	Claim 9; Page; 69pp; English.		
XX			
XX			
CC	The invention relates to nucleic acid molecules comprising altered HCV		
CC	NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)		
CC	internal ribosome entry site (IRES) region coding for one or more NS3,		
CC	NS5A, or EMCV IRES mutations, respectively. The location of the mutations		
CC	are detailed in the specification. Also included are (1) an expression		
CC	vector comprising a nucleotide sequence coding for the altered nucleic		
CC	acids, which is transcriptionally coupled to an exogenous promoter; (2) a		
CC	recombinant cell human hepatoma cell comprising the altered nucleic acids		
CC	; (3) a recombinant cell produced by introducing into a human hepatoma		
CC	cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)		
CC	replicon enhanced cell or which containing a functional HCV replicon; (5)		
CC	an HCV replicon enhanced cells made in the method; and (6) measuring the		
CC	ability of a compound to affect HCV activity. The HCV replicons and HCV		
CC	replicon enhanced cells are useful in studying HCV replication and		
CC	expression, and HCV and host cell interactions, producing HCV RNA and		
CC	proteins, and providing a system for measuring the ability of a compound		
CC	to modulate one or more HCV activities e.g. to discover drugs which may		
CC	treat HCV mediated diseases such as liver failure, cirrhosis and		
CC	hepatocellular carcinoma. The present sequence is an HCV replicon Con 1		
CC	mutant of the invention. Note: The present sequence is not shown in the		

CC specification but was created by the indexer using the HCV sequence
 CC appearing as ABK91411 and the information in Claim 9

XX Sequence 9605 BP; 1909 A; 2883 C; 2734 G; 2079 T; 0 U; 0 Other;

Query Match 86.9%; Score 1572.2; DB 6; Length 9605;

Best Local Similarity 91.8%; Pred. No. 0;

Matches 1661; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

```

QY 1 ATGAGCAGCAATCTTAACCTCAAGAAAAAACCAACGTAACCAACCGCGCCCAACAG 60
DB 342 ATGAGCAGCAATCTTAACCTCAAGAAAAAACCAACGTAACCAACCGCGCCCAACAG 401
QY 61 GAGGTCAATTCCTCCGCGCGGTGTGATGATGTTGTGTGATTTACTGTTCCTGCGCCAG 120
DB 402 GAGGTCAATTCCTCCGCGCGGTGTGATGATGTTGTGTGATTTACTGTTCCTGCGCCAG 461
QY 121 GAGCCAGGTGGGTGTGCGCGCACTAGAGAACTTCGAGCGGTGCAACCTCGTGA 180
DB 462 GAGCCAGGTGGGTGTGCGCGCACTAGAGAACTTCGAGCGGTGCAACCTCGTGA 521
QY 181 AGGCGCAACCTATCCCAAGGCTGCGCAAGCCGAGGCGAGGCTTGGGCTCAAGCCCG 240
DB 522 AGGCGCAACCTATCCCAAGGCTGCGCAAGCCGAGGCGAGGCTTGGGCTCAAGCCCG 581
QY 241 TACCTTGGGCGCTTATGAGCAAGAGGATGGGGTGGGAGATGCTCTGTCACCC 300
DB 582 TACCTTGGGCGCTTATGAGCAAGAGGATGGGGTGGGAGATGCTCTGTCACCC 641
QY 301 CGTGGCTCCGCGCTGATGGGAGCCCACTGACCCCGCGGTAGTGCCTAATTTGGT 360
DB 642 CGTGGCTCCGCGCTGATGGGAGCCCACTGACCCCGCGGTAGTGCCTAATTTGGT 701
QY 361 AAGTCATGCAATCCCTCAAGCTGCGGCTTGCAGCTCATGGGGTATCATTCGCTGTC 420
DB 702 AAGTCATGCAATCCCTCAAGCTGCGGCTTGCAGCTCATGGGGTATCATTCGCTGTC 761
QY 421 GGGGCTCCCTTAGGGGGCGCTGCGAGGCGCTTGGCGCATGGCGGTTCGAGAGAC 480
DB 762 GGGGCTCCCTTAGGGGGCGCTGCGAGGCGCTTGGCGCATGGCGGTTCGAGAGAC 821
QY 481 GGGGTAACCTATGCAAGGAAATTTAACCGGTTGCTCTTCTCTATCTTCTCTGAGCT 540
DB 822 GGGGTAACCTATGCAAGGAAATTTAACCGGTTGCTCTTCTCTATCTTCTCTGAGCT 881
QY 541 TTGCTGTCTGTTGAACATTCAGCTTCGCTTATGAGTGGCAACGTCGCGGATC 600
DB 882 TTGCTGTCTGTTGAACATTCAGCTTCGCTTATGAGTGGCAACGTCGCGGATC 941
QY 601 TACCATGTCAAGAACGATGCTCCAACTCAAGCATGTGTAAGAGACAGGACATGATC 660
DB 942 TACCATGTCAAGAACGATGCTCCAACTCAAGCATGTGTAAGAGACAGGACATGATC 1001
QY 661 ATSCAACACCCCGGGGTGTGCTGCTGTGTCGCGGAGGGTAAATTCCTCCGCTGCGG 720
DB 1002 ATSCAACACCCCGGGGTGTGCTGCTGTGTCGCGGAGGGTAAATTCCTCCGCTGCGG 1061
QY 721 GGGCTCACTCCACGCTCGCGGCAAGAGCGCAGCATCCCACTGGCAATACAGCGC 780
DB 1062 GGGCTCACTCCACGCTCGCGGCAAGAGCGCAGCATCCCACTGGCAATACAGCGC 1121
QY 781 CAGGTGATTTGCTGTTGGGGGCGCTGCTTCTGCTCCGCTATGTAAGTGGGAGATCTC 840
DB 1122 CAGGTGATTTGCTGTTGGGGGCGCTGCTTCTGCTCCGCTATGTAAGTGGGAGATCTC 1181
QY 841 TGGGATCTGTTTCTCGCTCTCAGCTGTTCACTTTCGCGCTCGCCAGCATCAGACG 900
DB 1182 TGGGATCTGTTTCTCGCTCTCAGCTGTTCACTTTCGCGCTCGCCAGCATCAGACG 1241
QY 901 GTACAGACTGCAATGTTCAATCTATCCGCGCAAGCATCAGGCTCAACGCGATGGCTGG 960
DB 1242 GTACAGACTGCAATGTTCAATCTATCCGCGCAAGCATCAGGCTCAACGCGATGGCTGG 1301

```

```

QY 961 GATATGATGATGAAGTGTGCTACCTACAGACGCTTGTGTATGCGAGTACTCGGATC 1020
DB 1302 GATATGATGATGAAGTGTGCTACCTACAGACGCTTGTGTATGCGAGTACTCGGATC 1361
QY 1021 CCACAGCTGTGTGACATGTGTGCGGGGCGCCACTGGGAGATCTGCGGGCTTGGC 1080
DB 1362 CCACAGCTGTGTGACATGTGTGCGGGGCGCCACTGGGAGATCTGCGGGCTTGGC 1421
QY 1081 TACTATTCATGTGTGGGAACTGCGGTAAAGTCTTGTGTGATGCTACTTTCGCGCG 1140
DB 1422 TACTATTCATGTGTGGGAACTGCGGTAAAGTCTTGTGTGATGCTACTTTCGCGCG 1481
QY 1141 GTTGAAGGGAACTTACACAGACAGGGGGAACACAGCGCGCGCCACGAGCTTACA 1200
DB 1482 GTTGAAGGGAACTTACAGAGAGGGGGAACAGATGGCCAAAAACACCTCGGATTAAG 1541
QY 1201 TCCCTCTTACACCTTGGGCGGCTCAGAAATTCAGCTTGTAAACCAACGCGACCTGG 1260
DB 1542 TCCCTCTTACACCTTGGGCGGCTCAGAAATTCAGCTTGTAAACCAACGCGACCTGG 1601
QY 1261 CACATCAACAGAACTGCTTGAATGCAATGACTCCCTCCAAACTGGGTTCTTGGCGCG 1320
DB 1602 CACATCAACAGAACTGCTTGAATGCAATGACTCCCTCCAAACTGGGTTCTTGGCGCG 1661
QY 1321 CTGTTCTACAGCAGATTCAATGCGTCCGAGTGTCTAGAGGCGATGGCCAGCTGCCG 1380
DB 1662 CTGTTCTACAGCAGATTCAATGCGTCCGAGTGTCTAGAGGCGATGGCCAGCTGCCG 1721
QY 1381 CCCATTGACAGTTCGATCAGGAGGTGGGCTCCCATCATCTTAATATGATCCCAAGCTTG 1440
DB 1722 CCCATTGACAGTTCGATCAGGAGGTGGGCTCCCATCATCTTAATATGATCCCAAGCTTG 1781
QY 1441 GACCAAGGCGCTTATGCTGCGCACTACGCACTCAACCGTGTGTATGTCGCGCGCTTG 1500
DB 1782 GACCAAGGCGCTTATGCTGCGCACTACGCACTCAACCGCGGTATGTCGCGCGCTTG 1841
QY 1501 CAGGTGTGGGCGCAGTGTACGTTTCACTCCAAAGCCGTTGTGTGGGAGACAGCGAT 1560
DB 1842 CAGGTGTGGGCGCAGTGTACGTTTCACTCCAAAGCCGTTGTGTGGGAGACAGCGAT 1901
QY 1561 CGTTTGGGCGCCCTACGTAAGATGGGTTGAGATGAGACGACGCTGCTTCTCAAC 1620
DB 1902 CGTTTGGGCGCCCTACGTAAGATGGGTTGAGATGAGACGACGCTGCTTCTCAAC 1961
QY 1621 AACACGCGCGCGCACGCGGCAACTGCTTGTGCTGTAATGATGATATGACACCGGTTTC 1680
DB 1962 AACACGCGCGCGCACGCGGCAACTGCTTGTGCTGTAATGATGATATGACACCGGTTTC 2021
QY 1681 ACCAAGAGTGTGGGGGCGCCCGCTGCAACATCGGGGGGTGGGCAACACTTGTATC 1740
DB 2022 ACCAAGAGTGTGGGGGCGCCCGCTGTAACATCGGGGGGTGGGCAACACTTGTATC 2081
QY 1741 TGCCCAAGCACTGCTTCCGGAAGCATCCGAGGCCACTTAAACCAATGCGGTTGGGG 1800
DB 2082 TGCCCAAGCACTGCTTCCGGAAGCATCCGAGGCCACTTAAACCAAGTGTGTTGGGG 2141
QY 1801 CCTTGGTTG 1809
DB 2142 CCTTGGTTG 2150

```

RESULT 10

ABK91426 standard; DNA; 9605 BP.

ABK91426;

15-NOV-2002 (first entry)

Hepatitis C virus Con 1 isolate DNA mutant 3.
 HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;
 hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;

KW internal ribosome entry site; IRES; NS5A; HCV replication; mutant.
XX Hepatitis C virus.
OS Synthetic.
XX
XX Key Location/Qualifiers
XX CDS 342..9374
XX /tag= a
XX /product= "HCV polyprotein"
XX /note= "The polyprotein consists of the Core, E1, E2, P7,
XX NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
XX mutation replace(4380,G)
XX /tag= b
XX WO200259321-A2.
XX 01-AUG-2002.
XX 16-JAN-2002; 2002WO-EP000526.
XX 23-JAN-2001; 2001US-0263479P.
XX (RICE-) IST RICERHE BIOL MOLECOLARE ANGELETTI.
XX De Francesco R, Migliaccio G, Paonessa G;
XX WPI; 2002-599793/64.
XX
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
XX NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
XX ribosome entry site (IRES) region, useful in studying HCV replication and
XX expression.
XX
XX Claim 9; Page: 69pp; English.
XX
XX The invention relates to nucleic acid molecules comprising altered HCV
XX NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
XX internal ribosome entry site (IRES) region coding for one or more NS3,
XX NS5A, or EMCV IRES mutations, respectively. The location of the mutations
XX are detailed in the specification. Also included are (1) an expression
XX vector comprising a nucleotide sequence coding for the altered nucleic
XX acids, which is transcriptionally coupled to an exogenous promoter; (2) a
XX recombinant cell human hepatoma cell comprising the altered nucleic acids
XX; (3) a recombinant cell produced by introducing the altered nucleic acids
XX cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
XX replicon enhanced cell or which containing a functional HCV replicon; (5)
XX an HCV replicon enhanced cells made in the method; and (6) measuring the
XX ability of a compound to affect HCV activity. The HCV replicons and HCV
XX expression, and HCV and host cell interactions, producing HCV RNA and
XX proteins, and providing a system for measuring the ability of a compound
XX to modulate one or more HCV activities e.g. to discover drugs which may
XX treat HCV mediated diseases such as liver failure, cirrhosis and
XX hepatocellular carcinoma. The present sequence is an HCV replicon Con 1
XX mutant of the invention. Note: The present sequence is not shown in the
XX specification but was created by the indexer using the HCV sequence
XX appearing as ABK91411 and the information in Claim 9
XX
XX Sequence 9605 BP; 1911 A; 2883 C; 2732 G; 2079 T; 0 U; 0 Other;
SQ
Query Match 86.9%; Score 1572.2; DB 6; Length 9605;
Best Local Similarity 91.8%; Pred. No. 0;
Matches 1661; Conservative 0; Mismatches 148; Indels 0; Gaps 0;
QY 1 ATGACGACGATCTCTTAACCTCAAGAAAAACCAACGTAACCAACGCGCGCCACAG 60
DB 342 ATGACGACGACGATCTTAACCTCAAGAAAAACCAACGTAACCAACGCGCGCCACAG 401
QY 61 GAGCTCAAGTTCGCGGCGGATGATCGATCGTGTGAGATTACCTGTTGCGGACAG 120
DB 402 GAGCTCAAGTTCGCGGCGGATGATCGATCGTGTGAGATTACCTGTTGCGGACAG 461
QY 121 GAGCCAGAGTTGGGTGTGCGCGACGTAAGAAAGATTCCGAGCGGTGCGAACCTCGTGA 180

DB 462 GAGCCAGAGTTGGGTGTGCGCGACGTAAGAAAGATTCCGAGCGGTGCGAACCTCGTGA 521
QY 181 AGGCGACAACTATTCCTCCCAAGGCTCCGACGCGGACAGGCGCTGGGCTCAAGCCGCG 240
DB 522 AGGCGACAACTATTCCTCCCAAGGCTCCGACGCGGACAGGCGCTGGGCTCAAGCCGCG 581
QY 241 TACCTTGGCCCTTATATGACAGAGGCAATGGGATGGGAGATGGCTCTCTGACCC 300
DB 582 TACCTTGGCCCTTATATGACAGAGGCTTGGGATGGGAGATGGCTCTCTGACCC 641
QY 301 GGTGCTCCGCGCTAGTTGGGCGCCCACTGACCCCGCGCGTAGGCTGCTAATTTGGGT 360
DB 642 GTGAGCTCTCGGCTAGTTGGGCGCCCACTGACCCCGCGCGTAGGCTGCTAATTTGGGT 701
QY 361 AAGTCATCGATACCTTCATATGAGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 420
DB 702 AAGTCATCGATACCTTCATATGAGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 761
QY 421 GAGCGCTCCCTTATGAGGCGGCTGCGAGGCGCTGCGGCTGCGGCTTCTGAGAGAC 480
DB 762 GAGCGCTCCCTTATGAGGCGGCTGCGAGGCGCTGCGGCTGCGGCTTCTGAGAGAC 821
QY 481 GAGCGTGAATATGCAACAGGGAATTTACCGGCTGCTTCTTCTATCTTCTTGGCT 540
DB 822 GAGCGTGAATATGCAACAGGGAATTTACCGGCTGCTTCTTCTATCTTCTTGGCT 881
QY 541 TTGCTGTCCTGTTTATGACATTCAGCTTCGCTTATGAGGCGGCAACGTCGCGGATC 600
DB 882 TTGCTGTCCTGTTTATGACATTCAGCTTCGCTTATGAGGCGGCAACGTCGCGGATC 941
QY 601 TACCATGTCAAGAAAGATGCTCCCACTCAAGCATGCTGTATGAGAGACGCGAGATGATC 660
DB 942 TACCATGTCAAGAAAGATGCTCCCACTCAAGCATGCTGTATGAGAGAGCGAGATGATC 1001
QY 661 ATGACACCCCGGAGTGTGTCCTGTCGCGAGGGAATTCCTCCGCTGCTGGTA 720
DB 1002 ATGACATACCCCGGAGTGTGTCCTGTCGCGAGGGAATTCCTCCGCTGCTGGTA 1061
QY 721 GCGCTACCTCCGAGCTGCGGCGGCAAGAGCGCGAGCATCCCACTGCGAGAAATGACGCG 780
DB 1062 GCGCTACCTCCGAGCTGCGGCGGCAAGAGCGCGAGCATCCCACTGCGAGAAATGACGCG 1121
QY 781 CAGCTGATTTGCTGCTGTTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
DB 1122 CAGCTGATTTGCTGCTGTTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1181
QY 841 TCGGATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
DB 1182 TCGGATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1241
QY 901 GTACAGGATGCAATTTGTTCAATCTATCCGCGGCAAGATGAGTCAACCGATGGCTGG 960
DB 1242 GTACAGGATGCAATTTGTTCAATCTATCCGCGGCAAGATGAGTCAACCGATGGCTGG 1301
QY 961 GATATGATGATGAATGCTGCTCACTTACAGACGCTTATGATGCAAGCTTCTCGGATC 1020
DB 1302 GATATGATGATGAATGCTGCTCACTTACAGACGCTTATGATGCAAGCTTCTCGGATC 1361
QY 1021 CCAGAGCTGTGAGCATGTGCGCGGCGGCACTGAGGAGTCTGCGGCGCTTGGC 1080
DB 1362 CCAGAGCTGTGAGCATGTGCGCGGCGGCACTGAGGAGTCTGCGGCGCTTGGC 1421
QY 1081 TACTATTCATGATGAGGGAATGAGGCTTATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1422 TACTATTCATGATGAGGGAATGAGGCTTATGATGATGATGATGATGATGATGATGATGAT 1481
QY 1141 GTTACAGGGAACCTTACACAGAGGGAACACAGCGCGCGCGCTTACCA 1200
DB 1482 GTTACAGGGAACCTTACAGAGGGAACAGGGAACAGTGGCAAAAACACCTCGGATTAAG 1541
QY 1201 TCCCTTTACACCTGCGGCGGCTGAGAAATCAAGCTTGAACACCAACGAGGAGCTGG 1260

Db	1542	TCCTCTTTTACCCCGGGTCAATCCAGAAATTCAGCTTGTAAACACCAAGGACGTGG	1601
Qy	1261	CACATCAACAGAACTGCTCTTGAAGTCAGTACTCCCTCCAACTGGGTTCTTGGCCGG	1320
Db	1602	CACATCAACAGAACTGCTCTTGAAGTCAGTACTCCCTCCAACTGGGTTCTTGGCTGG	1661
Qy	1321	CTGTTCTAACAGCAAGGTTCAATGCGTCCGGAATGCTCAAGGCCATGACCAGCTGCCG	1380
Db	1662	CTGTTCTAACAGCAAGGTTCAACTCACTTGTGAATGCCAGAGCCATGCGCATGCTGACG	1721
Qy	1381	CCCATTTGACCAAGTTCAGATTCAGAGGGTGGGGGTCCCATCACTTATTAATAGTCCCAAGGCTTG	1440
Db	1722	CCCATTCAGACGGGTTTCGCTCAGAGGGGTGGGGGCCCATCACTTATTAATAGTCAACAGCTCG	1781
Qy	1441	GACCAAGAGGCCCTTATTTGCTGGCACTACGACCTTCACCGTGTGATCTGTCGCCGCTTG	1500
Db	1782	GACCAAGAGGCCCTTATTTGTTGGCACTACGCAACCCGGGCGGTGCGGATGTATCCCGCGAGCG	1841
Qy	1501	CAGGTGTGTGGCCCAAGTATCTGTTTCACTCTCAAGCCCTGTTGTGTGGGGACGACCGGAT	1560
Db	1842	CAGGTGTGTGGTTCAGTGTACTGCTTCACCCCAAGCCCTGTCTGTGTGGGGACGACCGGAC	1901
Qy	1561	CGTTTCGGCGGCCCTTACGATCAGATGTGGGGTGAATGAGACGAGCGTGTCTTCTCAAC	1620
Db	1902	CGTTTCGGCGGCCCTTACGATCAGTGTGGGGGAGATGAGACGAGCGTGTCTTCTTAAAC	1961
Qy	1621	AACACGCGCGGCCCAACGAGGCACTGTTTCGCTGTACATGATGAATAGACACCGGGTTC	1680
Db	1962	AACACGCGCGGCCCGGCAAGGCACTGTTTGGCTGTACATGATGAATAGACACTGGGTTTC	2021
Qy	1681	ACGAAGACGTGTGGGGGCCCCCGGTGCAACATCTGGGGGGGTGGCGAACACACTTTGATC	1740
Db	2022	ACGAAGACGTGTGGGGGGCCCCCGGTGTAAACATCTGGGGGGATCTGGCAATAAACCTTGACC	2081
Qy	1741	TGCCCCCAGCACTGCTTCCGGAGAGATCCGAGGCGACTTTCACACCAATGCGGTTCCGGGG	1800
Db	2082	TGCCCCCAGCACTGCTTCCGGAGAGACCCCGAGGCACTTTCACCAAGTGTGTGTGGGG	2141
Qy	1801	CCTTGGTTTG 1809	
Db	2142	CCTTGGTTTG 2150	
RESULT 11			
ABK91433			
ID	ABK91433	standard; DNA; 9605 BP.	
XX	AC	ABK91433;	
XX	XX		
XX	DT	15-NOV-2002. (first entry)	
XX	XX		
DE	XX	Hepatitis C virus Con 1 isolate DNA mutant 10.	
KM	XX	HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;	
KW	KM	hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;	
KW	KW	internal ribosome entry site; IRES; NS5A; HCV replication; mutant.	
XX	XX		
OS	OS	Hepatitis C virus.	
XX	XX	Synthetic.	
XX	XX		
XX	Key	Location/Qualifiers	
XX	XX	342..9374	
XX	FT	/*tag= a	
XX	FT	/product= "HCV polypeptide"	
XX	FT	/note= "The polypeptide consists of the Core, E1, E2, P7,	
XX	FT	NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"	
XX	FT	mutation	
XX	FT	replase (6953..C)	
XX	XX	/*tag= b	
XX	XX		
XX	PD	01-AUG-2002.	

PF	16-JAN-2002; 2002MCO-EP000526.
PR	XX
PR	XX
XX	23-JAN-2001; 2001US-0263479P.
PA	(RICE-) IST RICERCHE BIOL MOLECOLARE ANGELTETTI.
PI	De Francesco R, Miglincio G, Paonessa G;
XX	WPI; 2002-599793/64.
DR	XX
XX	New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
PT	NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
PT	ribosome entry site (IRES) region, useful in studying HCV replication and
PT	expression.
XX	
PS	Claim 9; Page; 69pp; English.
CC	The invention relates to nucleic acid molecules comprising altered HCV
CC	NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
CC	internal ribosome entry site (IRES) region coding for one or more NS3,
CC	NS5a, or EMCV IRES mutations, respectively. The location of the mutations
CC	are detailed in the specification. Also included are (1) an expression
CC	vector comprising a nucleotide sequence coding for the altered nucleic
CC	acids, which is transcriptionally coupled to an exogenous promoter; (2) a
CC	recombinant cell human hepatoma cell comprising the altered nucleic acids
CC	; (3) a recombinant cell produced by introducing into a human hepatoma
CC	cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
CC	replicon enhanced cell or which containing a functional HCV replicon; (5)
CC	HCV replicon enhanced cells made in the method; and (6) measuring the
CC	ability of a compound to affect HCV activity. The HCV replicons and HCV
CC	replicon enhanced cells are useful in studying HCV replication and
CC	expression, and HCV and host cell interactions, producing HCV RNA and
CC	proteins, and providing a system for measuring the ability of a compound
CC	to modulate one or more HCV activities e.g. to discover drugs which may
CC	treat HCV mediated diseases such as liver failure, cirrhosis and
CC	hepatocellular carcinoma. The present sequence is an HCV replicon Con 1
CC	mutant of the invention. Note: The present sequence is not shown in the
CC	specification but was created by the index user using the HCV sequence
CC	appearing as ABK91411 and the information in Claim 9
XX	
SQ	Sequence 9605 BP; 1910 A; 2882 C; 2733 G; 2079 T; 0 U; 1 Other;
	Query Match 86.9%; Score 1572.2; DB 6; Length 9605;
	Best Local Similarity 91.8%; Pred. No. 0;
	Matches 1661; Conservative 0; Mismatches 148; Indels 0; Gaps 0;
OY	1 ATGAGCAGCATCTCTTAACCTCAAGAAGAAAACAACGTATACCAACACCGGCCACAG 60
DB	342 ATGAGCAGCATCTCTTAACCTCAAGAAGAAAACAACGTATACCAACACCGGCCACAG 401
OY	61 GACGTCAAAGTTCCCGGGGCGTGTCAGATCGTTGGTGAGTTTACTTTCGCCGCAGG 120
DB	402 GACGTCAAAGTTCCCGGGGCGTGTCAGATCGTTGGTGAGTTTACTTTCGCCGCAGG 461
OY	121 GGCCCCAAGTTGGGATGTGCGCGGCACTTAGAAGAACTTTCGAACGAGTCCGAACCTTG 180
DB	462 GGCCCCAAGTTGGGATGTGCGCGGCACTTAGAAGAACTTTCGAACGAGTCCGAACCTTG 521
OY	181 AGGGGACAACCTATATCCCAGAGGCTTCGACGCCGAGGGCAGAGGCTCGGACTCAAGCCCGG 240
DB	522 AGGGGACAACCTATATCCCAGAGGCTTCGACGCCGAGGGTAAAGGGCTCAAGCCCGG 581
OY	241 TACCCTTGGCCCTCTATATGGACAAGAGGGATAGGGGATGGAGAGTGGCTCTGTCAACC 300
DB	582 TACCCTTGGCCCTCTATATGGACAAGAGGGATAGGGGATGGAGAGTGGCTCTGTCAACC 641
OY	301 CGTGAGCTCCGCGCTAGTTGGGAGCCCACTGACCCCGGCGTAGGTGCGATATTGGGT 360
DB	642 CGTGAGCTTCGGCGCTAGTTGGGAGCCCACTGACCCCGGCGTAGGTGCGCAATTGGGT 701
OY	361 AAAGTCATCATACCTCTCACATGTGGGCTTCCCGCACTCATGGGGTACATTCGGCTGTC 420
DB	702 AAAGTCATCATACCTCTCACATGTGGGCTTCCCGCACTCATGGGGTACATTCGGCTGTC 761

421 GCGCTCCCTTAGGGGGCGCTGCGAGGGCCCTGCGCATGCGCTCGGGTTCTGAGAGAC 480
762 GGGGCCCCCTTAGGGGGCGCTGCGAGGGCCCTGCGCATGCGCTCGGGTTCTGAGAGAC 821
481 GCGCTGAATCTATGCAACAGGGAAATTTACCGGTTGCTCTTCTATCTCTTCTTGGCT 540
822 GCGCTGAATCTATGCAACAGGGAAATTTACCGGTTGCTCTTCTATCTCTTCTTGGCT 881
541 TTGCTGCTGTTTGAACATTCAGCTTCGGCTTATGAAGTGGCAAGCTGCGGGATC 600
882 TTGCTGCTGTTTGAACATTCAGCTTCGGCTTATGAAGTGGCAAGCTGCGGGATC 941
601 TACCATGTACAGAACGATGCTCAACTCAAGCATGCTGTAAGAGACAGAGCAATGATC 660
942 TACCATGTACAGAACGATGCTCAACTCAAGCATGCTGTAAGAGAGCAAGCAATGATC 1001
661 ATGCAACACCCCGGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
1002 ATGCAATACCCCGGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1061
721 GGGCTACCTCCAGCGCTGCGGGCAAGAGCGCAGATCCCACTGCGCAATATGAGAGC 780
1062 GGGCTACCTCCAGCGCTGCGGGCAAGAGCGCAGATCCCACTGCGCAATATGAGAGC 1121
781 CAGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
1122 CAGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1181
841 TCGCGATCTGTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
1182 TCGCGATCTGTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1241
901 GTACAGAGCTGAATTTGCTCAATCTATCCGGGCAAGTACAGGTACCGGAGCTGG 960
1242 GTACAGAGCTGAATTTGCTCAATCTATCCGGGCAAGTACAGGTACCGGAGCTGG 1301
961 GATATGATGATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
1302 GATATGATGATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1361
1021 CCACAAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
1362 CCACAAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1421
1081 TACTATTCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
1422 TACTATTCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1481
1141 GTTGAAGGGGAACTTACAGAGAGGGGAGACAGCGCGCGCGCGCGCGCGCGCGCTTACA 1200
1482 GTTGAAGGGGAACTTACAGAGAGGGGAGACAGCGCGCGCGCGCGCGCGCGCGCTTACG 1541
1201 TCCCTTTTCAACCTGCGCGCGCTGCAAAATTCAGCTTGTAAACCAACAGCGAGCTGG 1260
1542 TCCCTTTTCAACCTGCGCGCGCTGCAAAATTCAGCTTGTAAACCAACAGCGAGCTGG 1601
1261 CACATCAACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
1602 CACATCAACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1661
1321 CTGTTCTACAGCAGGTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
1662 CTGTTCTACAGCAGGTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1721
1381 CCCATTTGACAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
1722 CCCATTTGACAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1781
1441 GACGAGAGGCTTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
1782 GACGAGAGGCTTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1841

QY 1501 CAGGTGTGTGGCCAGATGATCTGTTTCACTCCAAAGCCCTGTGTGTGTGGAGACGACCGAT 1560
DB CAGGTGTGTGTGGCCAGATGATCTGTTTCACTCCAAAGCCCTGTGTGTGTGGAGACGACCGAT 1901
QY 1561 CGTTTGGGCGCCCTTACGTAAGATGGGTGAAGATGAGACGACGCTGCTTCTCAAC 1620
DB CGTTTGGGCGCCCTTACGTAAGATGGGTGAAGATGAGACGACGCTGCTTCTCAAC 1961
QY 1621 AACACGGGGCCGCAAGGGGCAACTGTTTGGGCTGATCATGATGATGATGACCGGGGTTG 1680
DB AACACGGGGCCGCAAGGGGCAACTGTTTGGGCTGATCATGATGATGATGACCGGGGTTG 2021
QY 1681 ACCAAGAGTGTGGGGGCCCCCGTGAACATCGGGGGGTGCGCAACACATTTGATC 1740
DB ACCAAGAGTGTGGGGGCCCCCGTGAACATCGGGGGGTGCGCAACATTTGATC 2081
QY 1741 TGCCCAACGACTGTTTCCGGAAGCATCCGAGGCACTTACACCAATGCGGTTTGGGG 1800
DB TGCCCAACGACTGTTTCCGGAAGCATCCGAGGCACTTACACCAATGCGGTTTGGGG 2141
QY 1801 CTTGGTTG 1809
DB CTTGGTTG 2150

RESULT 12
AAD25332
ID AAD25332 standard; cDNA; 9605 BP.
XX AAD25332;
XX
XX 12-MAR-2002 (first entry)
XX
XX Hepatitis C virus (HCV) full-length cDNA mutant #2.
DE
XX Hepatitis C virus; HCV; transfection; infection; virus neutralisation;
XX gene therapy; vaccine; immunoprotection; hepatotropic; viraucide; liver;
XX mutant; ss.
OS
XX Hepatitis C virus.
OS Synthetic.
XX
XX WO200189364-A2.
PN
XX
XX 29-NOV-2001.
PD
XX
XX 23-MAY-2001; 2001WO-US016822.
PF
XX
XX 23-MAY-2000; 2000US-00576989.
PR
XX
XX (UNITW) UNIV WASHINGTON.
PA
XX
XX Rice CM, Blight KJ;
PI
XX
XX WPI; 2002-066755/09.
DR
XX
XX
PT Hepatitis C virus variants having greater transfection efficiency and
PT ability to survive subpassage, useful as a vaccine for immunizing primate
PT to the virus, comprise non-naturally occurring viral sequences.
PS
XX
XX Example 2; Page 97-100; 174pp; English.
PS
XX
XX The invention relates to Hepatitis C virus (HCV) variants which include
XX polynucleotides comprising non-naturally occurring HCV sequence and HCV
XX variants that have a transfection efficiency and ability to survive
XX subpassage greater than HCV that have wild-type polypeptide coding
XX regions. The polynucleotides of the invention are useful for identifying
XX a cell line that is permissive for infection with HCV and detecting
XX replication of HCV in cells of the cell line. They are also useful for
XX testing a compound for anti-viral properties and for inhibiting HCV
XX infection. They are also useful for the generation of defined HCV virus
XX stocks to develop in vitro and in vivo assays for virus neutralisation,
XX attachment, penetration and entry, structure/function studies on HCV

CC proteins and RNA elements and identification of new antiviral targets, a
CC systematic survey of cell culture systems and conditions to identify
CC those that support wild-type and variant HCV RNA replication and particle
CC release, production of adaptive HCV variants capable of more efficiency
CC replication in cell culture, production of HCV variants with altered
CC tissue or species tropism, establishment of alternative animal models for
CC inhibitor evaluation including those supporting HCV variant replication,
CC development of cell-free HCV replication assays, production of
CC immunogenic HCV particles for vaccination, engineering of attenuated HCV
CC derivatives as possible vaccine candidates, engineering of attenuated or
CC defective HCV derivatives for expression of heterologous gene products
CC for gene therapy and vaccine applications and for utilization of the HCV
CC glycoproteins for targeted delivery of therapeutic agents to the liver
CC or other cell types with appropriate receptors. Vaccine comprising these
CC sequences is useful for inducing immunoprotection to HCV in a primate.
CC The present sequence is Hepatitis C virus (HCV) full-length cDNA
CC containing the mutation that results in Ser to Ile at position 1179 of
CC HCVrepBartman protein

Sequence 9605 BP; 1910 A; 2884 C; 2731 G; 2080 T; 0 U; 0 Other;

Query Match 86.9%; Score 1572.2; DB 6; Length 9605;

Best Local Similarity 91.8%; Pred. No. 0;

Matches 1661; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY 1 ATGAGCAGCAATCTAAACCTCAAGAAAAACCAACGTACACCAACGCGCGCCACAG 60
DB 342 ATGAGCAGCAATCTAAACCTCAAGAAAAACCAACGTACACCAACGCGCGCCACAG 401
QY 61 GAGGTCAAGTTCGCGGCGGTGTGTCAGATGTTGTGAGTTTACCTGTGCCGCGAG 120
DB 402 GAGGTCAAGTTCGCGGCGGTGTGTCAGATGTTGTGAGTTTACCTGTGCCGCGAG 461
QY 121 GGGCCAGGTGGGTGTGCGCGGAGTGAAGACTTCCGAGCGGTGCAACTCTGTGA 180
DB 462 GGGCCAGGTGGGTGTGCGCGGAGTGAAGACTTCCGAGCGGTGCAACTCTGTGA 521
QY 181 AGGCGCAACCTATCCCAAGGCTCCCGAGCGGAGGAGGAGGAGGAGGAGGAGGAGG 240
DB 522 AGGCGCAACCTATCCCAAGGCTCCCGAGCGGAGGAGGAGGAGGAGGAGGAGGAGG 581
QY 241 TACCTTGGCCCTCTATGAGCAAGAGGAGTGGGTCAGAGTGGCTCTGTCAACC 300
DB 582 TACCTTGGCCCTCTATGAGCAAGAGGAGTGGGTCAGAGTGGCTCTGTCAACC 641
QY 301 CGTGCTCCGCGCTATGTTGGGAGCCCACTGACCCCGCGGTAGTCCGCTAATTTGG 360
DB 642 CGTGCTCCGCGCTATGTTGGGAGCCCACTGACCCCGCGGTAGTCCGCTAATTTGG 701
QY 361 AAGATATGATACCTCTCATATGCGGCTTGCAGACTCATGGGATCATTCGCTGTC 420
DB 702 AAGATATGATACCTCTCATATGCGGCTTGCAGACTCATGGGATCATTCGCTGTC 761
QY 421 GGGGCTCCCTTAGGGGCGCTGCGAGGAGCTGGCGCATGGCGTCCGGGTTCTGAGGAC 480
DB 762 GGGGCTCCCTTAGGGGCGCTGCGAGGAGCTGGCGCATGGCGTCCGGGTTCTGAGGAC 821
QY 481 GGGGTAATATGCAACAGGAAATTTACCGGTTGCTTTCTATCTTCTCTTGGCT 540
DB 822 GGGGTAATATGCAACAGGAAATTTACCGGTTGCTTTCTATCTTCTCTTGGCT 881
QY 541 TTGCTGTCTGTGTTGACATTCAGGTTCCGCTTAAGAAATGGCAACGATCCGAGATC 600
DB 882 TTGCTGTCTGTGTTGACATTCAGGTTCCGCTTAAGAAATGGCAACGATCCGAGATC 941
QY 601 TACATATGACGAACATGTTGCTCAACTCAAGCATGTTGACGAGAGGAGCATGATC 660
DB 942 TACATATGACGAACATGTTGCTCAACTCAAGCATGTTGACGAGAGGAGCATGATC 1001
QY 661 ATGACACCCCGGGGTGTGTCCTGTGTCGCGGAGGATTAATCTTCCGCTGCTGGGTA 720
DB 1002 ATGACATACCCCGGGGTGTGTCCTGTGTCGCGGAGGATTAATCTTCCGCTGCTGGGTA 1061

QY 721 GGGCTACATCCACAGCTCGCGCAAGAGACGACGAGCATCCCACTGCGCAATACGACGC 780
DB 1062 GGGCTACATCCACAGCTCGCGCAAGAGACGACGAGCATCCCACTGCGCAATACGACGC 1121
QY 781 CAGGTGATTTGCTGTGGGGGCGGTGCTTCTGTGCTGCTATGATGATGGGGATTC 840
DB 1122 CAGGTGATTTGCTGTGGGGGCGGTGCTTCTGTGCTGCTATGATGATGGGGATTC 1181
QY 841 TGGGATCTGTTTCT 900
DB 1182 TGGGATCTGTTTCT 1241
QY 901 GATACAGATGCAATGTTCAATCTATCCCGGACGATTCAGAGTACCGCATGCTTGG 960
DB 1242 GATACAGATGCAATGTTCAATCTATCCCGGACGATTCAGAGTACCGCATGCTTGG 1301
QY 961 GATATGATGATGAATGTCACCTTACAGGACCTTATGATATGAGGATCTCCGATC 1020
DB 1302 GATATGATGATGAATGTCACCTTACAGGACCTTATGATATGAGGATCTCCGATC 1361
QY 1021 CCACAGCTGTGTGACATGTTGGGGGGGCGCCACTGGGAGTCTTGGCGGCTTGGC 1080
DB 1362 CCACAGCTGTGTGATATGATGATGATGATGATGATGATGATGATGATGATGATG 1421
QY 1081 TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1422 TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1481
QY 1141 GTTGAAGGGAACCTTACACGACAGGGGGGACACAGGCGCGCGCCACGCGCTTACA 1200
DB 1482 GTTGAAGGGAACCTTACACGACAGGGGGGACAGAGGCGCAAAACACCTCGGAGTAC 1541
QY 1201 TCCCTTTTCACTGCGCGCGCTCAGAAATTCAGCTTTTAAACCAACGCGACCTG 1260
DB 1542 TCCCTTTTCACTGCGCGCGCTCAGAAATTCAGCTTTTAAACCAACGCGACCTG 1601
QY 1261 CACATCAACAGATGCTTGAACGCAATGACTCCCTCAACACTGGGTTCTTGGCGCG 1320
DB 1602 CACATCAACAGATGCTTGAACGCAATGACTCCCTCAACACTGGGTTCTTGGCGCG 1661
QY 1321 CTGTTCTACACGACAGTTCATGATGATGATGATGATGATGATGATGATGATGATG 1380
DB 1662 CTGTTCTACACGACAGTTCATGATGATGATGATGATGATGATGATGATGATGATG 1721
QY 1381 CCCATGACAGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440
DB 1722 CCCATGACAGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1781
QY 1441 GACGAGAGGCGCTATGCTGAGCACTACGCAACGCTGTGTATGCTGTGCGCGGTTG 1500
DB 1782 GACGAGAGGCGCTATGCTGAGCACTACGCAACGCTGTGTATGCTGTGCGCGGTTG 1841
QY 1501 CAGGTGTGTGCGCAGTGTGATGTTTCACTCCAGACCTGTGTGTGTGGAGGACGACAT 1560
DB 1842 CAGGTGTGTGCGCAGTGTGATGTTTCACTCCAGACCTGTGTGTGTGGAGGACGACAT 1901
QY 1561 CGTTTGGCGCCCTTACGATAGATGGGTGAAATGAGACGACGATGCTTCTCAAC 1620
DB 1902 CGTTTGGCGCCCTTACGATAGATGGGTGAAATGAGACGACGATGCTTCTCAAC 1661
QY 1621 AACAGCGGCGCGCACGAGGCAATGTTCCGCTGATACATGATGATGATGATGATG 1680
DB 1962 AACAGCGGCGCGCACGAGGCAATGTTCCGCTGATACATGATGATGATGATGATG 2021
QY 1681 ACCAAGACGTGTGGGGGCGCCCGGTGCAACATCGGGGGGTGGGCAACACTTTGATC 1740
DB 2022 ACCAAGACGTGTGGGGGCGCCCGGTGCAACATCGGGGGGTGGGCAACACTTTGATC 2081
QY 1741 TGGCCCAAGGATGCTTCCGGAAGCATCCGAGGCACTTACACCAATGCGGTTGGGG 1800
DB 2082 TGGCCCAAGGATGCTTCCGGAAGCATCCGAGGCACTTACACCAATGCGGTTGGGG 2141
QY 1801 CCTTGGTTG 1809

```
|||||
|||||
|||||
DB      2142 CCTGTGTTG 2150

RESULT 13
ABK91427 standard; DNA; 9608 BP.
XX
AC ABK91427;
XX
XX 15-NOV-2002 (first entry)
XX
DE Hepatitis C virus Con 1 isolate DNA mutant 4.
XX
XX HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;
XX hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
XX internal ribosome entry site; IRES; NS5A; HCV replication; mutant.
XX
OS Hepatitis C virus.
XX
XX Synthetic.
XX
FH Key location/Qualifiers
FT CDS 342..9374
FT /tag= a
FT /product= "HCV polyprotein"
FT /note= "The polyprotein consists of the Core, E1, E2, P7,
FT NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
FT mutation
FT /tag= b

WO200259321-A2.
XX
XX 01-AUG-2002.
XX
XX 16-JAN-2002; 2002WO-BP000526.
XX
XX 23-JAN-2001; 2001US-0263479P.
XX
XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
XX
XX De Francesco R, Migliaccio G, Paonessa G;
XX
XX WPI; 2002-599793/64.
XX
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
XX NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
XX ribosome entry site (IRES) region, useful in studying HCV replication and
XX expression.
XX
XX Claim 9; Page; 69pp; English.
XX
XX The invention relates to nucleic acid molecules comprising altered HCV
XX NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
XX internal ribosome entry site (IRES) region coding for one or more NS3,
XX NS5A, or EMCV IRES mutations, respectively. The location of the mutations
XX are detailed in the specification. Also included are (1) an expression
XX vector comprising a nucleotide sequence coding for the altered nucleic
XX acids, which is transcriptionally coupled to an exogenous promoter; (2) a
XX recombinant cell human hepatoma cell comprising the altered nucleic acids
XX; (3) a recombinant cell produced by introducing into a human hepatoma
XX cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
XX replicon enhanced cell or which containing a functional HCV replicon; (5)
XX an HCV replicon enhanced cells made in the method; and (6) measuring the
XX ability of a compound to affect HCV activity. The HCV replicons and HCV
XX replicon enhanced cells are useful in studying HCV replication and
XX expression, and providing a system for measuring the ability of a compound
XX to modulate one or more HCV activities e.g. to discover drugs which may
XX treat HCV mediated diseases such as liver failure, cirrhosis and
XX hepatocellular carcinoma. The present sequence is an HCV replicon Con 1
XX mutant of the invention. Note: The present sequence is not shown in the
XX specification but was created by the indexer using the HCV sequence
XX appearing as ABK91411 and the information in Claim 9
```

```
XX
SQ      Sequence 9608 BP; 1913 A; 2883 C; 2733 G; 2079 T; 0 U; 0 Other;
Query Match      86.9%; Score 1572.2; DB 6; Length 9608;
Best Local Similarity 91.8%; Pred. No. 0;
Matches 1661; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

1 ATGAGCAGAACTCTAAACCTCAAGAAAAACCAAGTAAACACCAACCGCCGCCACAG 60
DB ATGAGCAGAACTCTAAACCTCAAGAAAAACCAAGTAAACACCAACCGCCGCCACAG 401
QY GACGTCAGATTCCCGGGGCGGTGTGATGATGCTGTGTGATGATGATGATGATGATG 120
DB GACGTCAGATTCCCGGGGCGGTGTGATGATGATGATGATGATGATGATGATGATG 461
QY GACCCCAAGTTCCCGGGGCGGTGTGATGATGATGATGATGATGATGATGATGATG 180
DB GACCCCAAGTTCCCGGGGCGGTGTGATGATGATGATGATGATGATGATGATGATG 521
QY AGGCGACAACTATCCCAAGGCTCGCAGCCCGAGGGCAGAGGCTGAGCTCAGCCGGG 240
DB AGGCGACAACTATCCCAAGGCTCGCAGCCCGAGGGTAGGGCTGAGCTCAGCCGGG 581
QY TACCTTGGCCCTCTATGCGAAGAGGCAATGGGGTGGGATGATGATGATGATGATG 300
DB TACCTTGGCCCTCTATGCGAAGAGGCAATGGGGTGGGATGATGATGATGATGATG 641
QY TACCCCTGGCCCTCTATGCGAAGAGGCAATGGGGTGGGATGATGATGATGATGATG 641
DB TACCCCTGGCCCTCTATGCGAAGAGGCAATGGGGTGGGATGATGATGATGATGATG 641
QY CGTGGCTCCCGGCTAGTGGGGGCGCCCACTGACCCCGGGGCTAGGTGGCGTAATTTGGGT 360
DB CGTGGCTCCCGGCTAGTGGGGGCGCCCACTGACCCCGGGGCTAGGTGGCGTAATTTGGGT 701
QY CGTGGCTCCCGGCTAGTGGGGGCGCCCACTGACCCCGGGGCTAGGTGGCGTAATTTGGGT 701
DB CGTGGCTCCCGGCTAGTGGGGGCGCCCACTGACCCCGGGGCTAGGTGGCGTAATTTGGGT 701
QY AAAGTCATGATACCTCATACGAGGCTTCCGCACTCATAGGGGTATCATTCGCTGCTC 420
DB AAAGTCATGATACCTCATACGAGGCTTCCGCACTCATAGGGGTATCATTCGCTGCTC 761
QY GCGCTCCCTTAAAGGGGCGCTGCGAGGGCCCTGGCGCATGGCGCTTCTGAGAGAC 480
DB GCGCTCCCTTAAAGGGGCGCTGCGAGGGCCCTGGCGCATGGCGCTTCTGAGAGAC 821
QY GCGCGCCCTTAAAGGGGCGCTGCGAGGGCCCTGGCGCATGGCGCTTCTGAGAGAC 821
DB GCGCGCCCTTAAAGGGGCGCTGCGAGGGCCCTGGCGCATGGCGCTTCTGAGAGAC 821
QY GCGCGTGAATGATGAGAGAGGAAATTTACCGGGTGTCTTCTGATGCTGCTGCTGCT 540
DB GCGCGTGAATGATGAGAGAGGAAATTTACCGGGTGTCTTCTGATGCTGCTGCTGCT 881
QY GCGCGTGAATGATGAGAGAGGAAATTTACCGGGTGTCTTCTGATGCTGCTGCTGCT 881
DB GCGCGTGAATGATGAGAGAGGAAATTTACCGGGTGTCTTCTGATGCTGCTGCTGCT 881
QY TTGCTGCTGCTTGTGACCATTCAGCTTCCGCTTATGAGTGGCGCAAGTGTCCGGAGTC 600
DB TTGCTGCTGCTTGTGACCATTCAGCTTCCGCTTATGAGTGGCGCAAGTGTCCGGAGTC 941
QY TTGCTGCTGCTTGTGACCATTCAGCTTCCGCTTATGAGTGGCGCAAGTGTCCGGAGTC 941
DB TTGCTGCTGCTTGTGACCATTCAGCTTCCGCTTATGAGTGGCGCAAGTGTCCGGAGTC 941
QY TACCATGTCAAGACGATGCTCCCAAGCAATGCTGTACGAGACAGCGGACATGATC 660
DB TACCATGTCAAGACGATGCTCCCAAGCAATGCTGTACGAGACAGCGGACATGATC 1001
QY ATGACACACCCCGGGGTGTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTC 720
DB ATGACACACCCCGGGGTGTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTC 1061
QY ATGACATACCCCGGGGTGTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTC 1061
DB ATGACATACCCCGGGGTGTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTC 1061
QY GCGCTCACTCCAGAGCTGCGCGGCAAGAGCCGACGATCCCACTGCGACAAATGACAGC 780
DB GCGCTCACTCCAGAGCTGCGCGGCAAGAGCCGACGATCCCACTGCGACAAATGACAGC 1121
QY GCGCTCACTCCAGAGCTGCGCGGCAAGAGCCGACGATCCCACTGCGACAAATGACAGC 780
DB GCGCTCACTCCAGAGCTGCGCGGCAAGAGCCGACGATCCCACTGCGACAAATGACAGC 1121
QY TGGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
DB TGGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1181
QY TGGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
DB TGGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1241
QY GTACAGAGCTCAATGCTCAATGCTCAATGCTCAATGCTCAATGCTCAATGCTCAATGCT 960
DB GTACAGAGCTCAATGCTCAATGCTCAATGCTCAATGCTCAATGCTCAATGCTCAATGCT 1301
QY GTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB GTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
```


QY	181	AGGCGACAACCTATATCCCAAGGCTGGCAGGCCGAGGGGACTGGGACTATAGCCCGGG	240
Db	1979	AGGCGACAACCTATATCCCAAGGCTGGCAGGCCGAGGGGAGGGGCTGGGACTATAGCCCGGG	2038
QY	241	TACCCCTGGCCCTCTATATGGACAAGAGGGCATGGGGGTGGGGCAGAGATGGCTCTGTACACC	300
Db	2039	TACCCCTGGCCCTCTATATGGACAATAGGGCTTTGGGGTGGGACAGATGGCTCTGTACACC	2098
QY	301	CGTGGCTCCCGGCTTAGTGTGGGGCCCACTGACCCCGGCGGTAGGTGCGTAATTTGGGT	360
Db	2099	CGTGGCTCTCCGGCTTAGTGTGGGGCCCAAGGACCCCGGCGTAAGGTTCGGCAATTTGGGT	2158
QY	361	AAAGTCATATGATACCTTCACATGCGGCTTCGCGGACTCTATATGGGGTACATTCGCTGCTC	420
Db	2159	AAAGTCATATGATACCTTCACATGCGGCTTCGCGGACTCTATATGGGGTACATTCGCTGCTC	2218
QY	421	GGCGCTCCCTTAGGGGGGGGCTGGCCAGAGGCCCTGGGGGAGATGGGGTCCGGGTTCTGGAGGAC	480
Db	2219	GGCGCCCCCTTAGGGGGGGGCTGGCCAGAGGGCCCTGGGGGAGATGGGGTTCGGAGGAGAC	2278
QY	481	GGCGGTAACTATGACAACAGGGGAATTTACCCGGTTGCTCTCTATCTTCCTCTTGGCT	540
Db	2279	GGCGGTAACTATGACAACAGGGGAATTTGCCCCGGTTGCTCTCTTATCTTCCTTTTGGCT	2338
QY	541	TTGCTGTCTCTGTTTGAACAATTCAGCTTCGCTTATGAATGGCGCAAGCTGTCCGGATTC	600
Db	2339	TTGCTGTCTCTGTTTGAACAATTCAGCTTCGCTTATGAATGGCGCAAGTATCCGGAGTGG	2398
QY	601	TACCATATGACAGAACGATTTGCTCCCAACTCAACATCGTGTACGAGACAGCGGACATATGATC	660
Db	2399	TACCATATGACAGAACGATTTGCTCCCAACTCAACATCGTGTATGAGAGCAGCGGACATATGATC	2458
QY	661	ATGCAACACCCCGGGGTGTGTGCTCTGTGTCCGGAGAGGTAATTTCTCCGCTGCTGGGTA	720
Db	2459	ATGCAATACCCCGGGGTGTGTGCTCTGTGTCCGGAGAGAACATCTCTCCGCTGCTGGGTA	2518
QY	721	GCGCTCACTCCCAAGCTGGCGGCCAAGAGCGCAGCATTCCTCACTGCAATATGACAGCGC	780
Db	2519	GCGCTCACTCCCAAGCTGGCGGCCAAGAGCGCTACGCTCCCACTACAGACGATATGACAGCGC	2578
QY	781	CACGTGCATTTGCTGTGTGGGGGGGCTGCTCTGTCTCCGCTATATGATAGTGGGGGATCTC	840
Db	2579	CATGTGTGATTTGCTGTGTGGGGGGGCTGTCTCTGTCTCCGCTATATGATAGTGGGGAATCTC	2638
QY	841	TGCGGATCTGTTTTCCTGCTCTCAGCGTGTACATTCCTGCGCTCGCGACATCAGACG	900
Db	2639	TGCGGATCTGTTTTCCTGCTCGCCAGACGTGTTCACCTTCTCGGCTTCGCGGACAGACACA	2698
QY	901	GTACAGACATGCAATTTGTTCAATCTATCCCGGCAAGTACAGATCACCGCATGGCTTGG	960
Db	2699	GTACAGACATGCAATTTGCTCAATATATCCCGGCAAGTACAGATCACCGTATGGCTTGG	2758
QY	961	GATATGATATGATGAATCTGTGTCACCTTACAGACGCCCTTAGTGTATGCGAGCTACTCCGATTC	1020
Db	2759	GATATGATATGATGAATCTGTGTCACCTTACAGACGCCCTTAGTGTATGCCAGTTACTCCGATTC	2818
QY	1021	CCACAAGCTGTGTGTGGAACATGGTGTGGGGGGGCGCACTGGGGAGTCTGTGGCGGGCTTGGC	1080
Db	2819	CCACAAGCTGTGTGTGATATGGTGTGGGGGGGCGCAATGGGGAGTCTTAGCGGGGCTTGGC	2878
QY	1081	TACTATTCATATGATGGGGGAACCTAGGCTAAGGCTTGTGTGTGTATCTACTCTTTTCCCGGC	1140
Db	2879	TACTATTCATATGATGGGGGAACCTAGGCTAAGGCTTGTGTGTGTATCTACTCTTTTCCCGGC	2938
QY	1141	GTTTACGGGGAACTTTTACAGACAGGGGGGGAACACAGGCGCGCGCCCAAGGGCTTAC	1200
Db	2939	GTTTACGGGGGAACCTATATGACAGGGGGGGAAGATGGGCCAATAAACCTCCGGGATTTACG	2998
QY	1201	TCCCTCTTACACTGTGGGCGCGGCTAGAAATTCAGCTTGTATACCAACAAAGGGCAGCTGG	1260
Db	2999	TCCCTCTTTTACCCGGGCTATTCAGAAATTCAGCTTGTATACCAACAAAGGGCAGCTGG	3058
QY	1261	CACATCAACAGATGTGCTTGAATGTGCATGATCTCCCTTCAAACTGGGGTTCCTTCCCGG	1320

Db	3059	CACATCAACAGACTGCCCTGAATCGAATGACTCCCTCAACACTGGGGTTCCTTGCTGC	3118
Qy	1321	CTGTATTACAGCAGAGTTCATATGCGTCCGATGCTCAGAGCGCATATGGCAGCTGC	1380
Db	3119	CTGTATTACAGCAGAGTTCATATGCGTCCGATGCTCAGAGCGCATATGGCAGCTGC	3178
Qy	1391	CCCATATGACAGTTTCGATCAGGGGTGGGTCCTCATCTATATATGATGCCAGGCTTG	1440
Db	3179	CCCATATGACAGTTTCGATCAGGGGTGGGTCCTCATCTATATGATGACACAGCTGC	3238
Qy	1441	GACCAAGAGCCCTTATGCTGCACTACGACCTCAACCGTGTGTATGTCGCCGCTTG	1500
Db	3239	GACCAAGAGCCCTTATGCTGCACTACGACCTCAACCGTGTGTATGTCGCCGCTTG	3298
Qy	1501	CAGGTGTGTGCGCCAGTGTATCTGTTTCACTTCAAGCCCTGTGTGTGTGGGAGCAGC	1560
Db	3299	CAGGT	3358
Qy	1561	CGTTTCGGCGCCCTTACGTAAGATGAGTGGAGTGAAGAGAGAGAGAGAGAGAGAG	1620
Db	3359	CGTTTCGGCGCCCTTACGTAAGATGAGTGGAGTGAAGAGAGAGAGAGAGAGAGAG	3418
Qy	1621	AACACGCGCGCCGCAAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1680
Db	3419	AACACGCGCGCCGCAAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3478
Qy	1681	ACCAAGAGT	1740
Db	3479	ACCAAGAGT	3538
Qy	1741	TGCCCCAGCAGTCTCTTCCGGAAGCATCCCGAGGCACTTACACCAATATCGGTTG	1800
Db	3539	TGCCCCAGCAGTCTCTTCCGGAAGCATCCCGAGGCACTTACACCAATATCGGTTG	3598
Qy	1801	CCTTGCTTG 1809	
Db	3599	CCTTGCTTG 3607	
RESULT 15			
AAA98965			
ID	AAA8965	standard; DNA; 11076 BP.	
XX	AC	AAA8965;	
XX	AA		
DT	08-FEB-2001	(first entry)	
XX	XX		
DE	Hepatitis C virus DNA fragment SEQ ID NO: 1.		
XX	XX		
KM	Cell culture; therapy; infection; vaccine; diagnosis; gene therapy; ds.		
XX	XX		
OS	Hepatitis C virus.		
XX	XX		
PN	DE19915178-A1.		
XX	PD		
XX	05-OCT-2000.		
PF	03-APR-1999;	99DE-01015178.	
XX	XX		
PR	03-APR-1999;	99DE-01015178.	
XX	XX		
PA	(UWMA-) UNIV MAINZ GUTENBERG JOHANNES.		
XX	XX		
FI	Bartenschlager R;		
XX	XX		
DR	WPI; 2000-629140/61.		
XX	XX		
PT	Cell culture system for hepatitis C virus, useful e.g. in screening for		
PT	therapeutic agents, comprises human hepatoma cells containing a viral RNA		
PT	construct that includes a selectable gene.		
XX	XX		
PS	Claim 5; Page 14-22; 58PP; German.		

Db 3613 CCTTGTTG 3621

Search completed: February 19, 2005, 01:26:50
Job time : 1018.71 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2005, 00:21:46 ; Search time 311.302 Seconds
(without alignments)
9508.540 Million cell updates/sec

Title: US-09-664-363-21_COPY_308_2116

Perfect score: 1809
Sequence: 1 ATGACGACCAATCTTAACCC.....GGCGTTGGGCGCTTGTTG 1809

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1809	100.0	2116	3	US-08-191-160-21 Sequence 21, Appl
2	1570.6	86.8	11076	4	US-09-539-601-1 Sequence 1, Appl
3	1570.6	86.8	11076	4	US-09-539-601-1 Sequence 19, Appl
4	1570.6	86.8	11076	4	US-09-539-601-25 Sequence 25, Appl
5	1570.6	86.8	11076	4	US-09-539-601-31 Sequence 31, Appl
6	1556.2	86.0	9595	3	US-09-014-416-4 Sequence 4, Appl
7	1556.2	86.0	9595	3	US-09-014-416-6 Sequence 6, Appl
8	1543.4	85.3	9413	4	US-09-827-688-6 Sequence 6, Appl
9	1538.6	85.1	2433	3	US-08-612-973-49 Sequence 49, Appl
10	1538.6	85.1	2433	3	US-08-612-973-49 Sequence 49, Appl
11	1531.2	84.6	9472	4	US-08-150-204E-96 Sequence 96, Appl
12	1517.8	83.9	6039	1	US-08-324-977-11 Sequence 11, Appl
13	1517.8	83.9	6039	2	US-08-324-977-11 Sequence 11, Appl
14	1517.8	83.9	6039	2	US-08-324-977-11 Sequence 11, Appl
15	1517.8	83.9	6039	3	US-08-904-686A-11 Sequence 11, Appl
16	1517.8	83.9	6039	3	US-09-315-850-11 Sequence 11, Appl
17	1517.8	83.9	9030	1	US-08-324-977-13 Sequence 13, Appl
18	1517.8	83.9	9030	2	US-08-324-977-13 Sequence 13, Appl
19	1517.8	83.9	9030	2	US-08-324-977-13 Sequence 13, Appl
20	1517.8	83.9	9030	3	US-08-324-977-13 Sequence 13, Appl
21	1517.8	83.9	9416	2	US-08-384-616-13 Sequence 13, Appl
22	1517.8	83.9	9416	2	US-08-384-616-13 Sequence 13, Appl
23	1517.8	83.9	9416	2	US-08-384-616-13 Sequence 13, Appl
24	1517.8	83.9	9416	2	US-08-384-616-13 Sequence 13, Appl
25	1517.8	83.9	9416	2	US-08-384-616-13 Sequence 13, Appl
26	1334.2	73.8	1539	2	US-08-823-895A-27 Sequence 27, Appl
27	1206.4	66.7	12980	3	US-08-470-426B-17 Sequence 14, Appl

28	1206.4	66.7	12980	3	US-09-034-756-5 Sequence 5, Appl
29	1204.8	66.6	9646	3	US-08-811-566-1 Sequence 1, Appl
30	1204.8	66.6	9646	3	US-09-034-756-1 Sequence 1, Appl
31	1202	66.4	2082	3	US-08-612-973-47 Sequence 47, Appl
32	1202	66.4	2082	3	US-08-927-597-47 Sequence 47, Appl
33	1201.6	66.4	9599	3	US-09-014-416-2 Sequence 2, Appl
34	1195.2	66.1	9365	4	US-08-827-688-7 Sequence 7, Appl
35	1195.2	66.1	9401	2	US-08-432-693-1 Sequence 1, Appl
36	1195.2	66.1	9416	3	US-08-811-566-19 Sequence 19, Appl
37	1195.2	66.1	9416	3	US-09-034-756-19 Sequence 19, Appl
38	1193.6	66.0	9416	3	US-08-823-895A-26 Sequence 26, Appl
39	1193.6	66.0	9416	4	US-10-104-966-13 Sequence 13, Appl
40	1180.8	65.3	9379	3	US-08-444-818-176 Sequence 176, App
41	1180.8	65.3	9401	5	PCT-US91-02225-9 Sequence 9, Appl
42	1177.6	65.1	8987	3	US-08-444-818-137 Sequence 137, App
43	1177.6	65.1	9185	3	US-08-444-818-122 Sequence 122, App
44	1177.6	65.1	9185	3	US-08-444-818-123 Sequence 123, App
45	1177.6	65.1	9379	3	US-09-388-874-1 Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-191-160-21
; Sequence 21, Application US/08191160
; Patent No. 6210675
; GENERAL INFORMATION:
; APPLICANT: Highfield, Peter Edmund
; APPLICANT: Rodgers, Brian Colin
; APPLICANT: Tedder, Richard Seton
; APPLICANT: Barbara, John Anthony James
; TITLE OF INVENTION: Viral Agent
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESSES:
; ADDRESSER: Rothwell, Figg, Ernst & Kurz
; STREET: 1700 K Street
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy diskette, 5.25 inch, 360 Kb storage
; OPERATING SYSTEM: IBM AT compatible
; SOFTWARE: Wordperfect 5.0 (DOS text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/191,160
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/628,516
; FILING DATE: 17 DEC 1990
; APPLICATION NUMBER: UK 89 28 562.1
; FILING DATE: 18 DEC 1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 90 04 414.0
; FILING DATE: 27 FEB 1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 90 04 814.1
; FILING DATE: 03 MAR 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: E. Anthony Figg
; REGISTRATION NUMBER: 27,195
; REFERENCE/DOCKET NUMBER: 1645-103A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 833-5740
; TELEFAX: (202) 833-5740
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2116 base pairs
; TYPE: nucleotide with corresponding protein
; STRANDEDNESS: single

```

1      TOPOLOGY: linear
2
3      MOLECULE TYPE: cDNA to genomic RNA
4
5      ORIGINAL SOURCE:
6
7      ORGANISM: human; serum infectious for PT-NANBH
8
9      IMMEDIATE SOURCE:
10
11     LIBRARY: cDNA clones from 5' end of the genome
12
13     FEATURE:
14
15     LOCATION: from 308 to 2116 bp start of the PT-NANBH
16
17     LOCATION: polyprotein
18
19     OTHER INFORMATION: viral structural and non-structural
20
21     OTHER INFORMATION: proteins
22
23     OS-08-191-160-21

```

Query Match	100.0%	Score 1809;	DB 3;	Length 2116;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1809; Conservative	0;	Mismatches	0;	Gaps 0;

OY	1	TTGAGACAGATTCCTTAAACCTTCAAGAAAAACCAAGTAACTACACCAACCCGCCCAACAG	60
Db	308	ATGAGACAGAAATCCTTAAACCTTCAAGAAAAACCAAGTAACTACCAACACCGCGCCCAAG	367
OY	61	GACGTCAGATTCCCGGCGGATGATCAGATTCGTTGGTAGATTTACCTGTGGCGCGCAGG	120
Db	368	GACGTCAGATTCCCGGCGGATGATCAGATTCGTTGGTAGATTTACTGTGGCGCGCAGG	427
OY	121	GGCCCAAGATTGAGTGTGTGCGCGCAGCTAGAAAGACTTCCGAGCGGTGCGAACCTTGTGGA	180
Db	428	GGCCCAAGATTGAGTGTGTGCGCGCAGCTAGAAAGACTTCCGAGCGGTGCGAACCTTGTGGA	487
OY	181	AGGCGACAACCTATCCCAAGGCTGCGCAGCCCGAGGGCAGGGGCTGGGACTCAGCCCGGG	240
Db	488	AGGCGACAACCTATCCCAAGGCTGCGCAGCCCGAGGGCAGGGGCTGGGACTCAGCCCGGG	547
OY	241	TACCCCTTGCCCTCTATATGCGACAAGAGGACATGAGGATGAGGCGAGATGAGCTCTGTCAACC	300
Db	548	TACCCCTTGCCCTCTATATGCGACAAGAGGACATGAGGATGAGGCGAGATGAGCTCTGTCAACC	607
OY	301	CGTGGCTCCCGGCTTAATGTTGGGGGCCCACTGACCCCGGCGTAAAGTCCGCTAATTTGGGT	360
Db	608	CGTGGCTCCCGGCTTAATGTTGGGGGCCCACTGACCCCGGCGTAAAGTCCGCTAATTTGGGT	667
OY	361	AAAGTCATCGATACCCCTACATGAGCGGGCTGCGCGACCTCATGAGGGTACATTCGCGTCGTC	420
Db	668	AAAGTCATCGATACCCCTACATGAGCGGGCTGCGCGACCTCATGAGGGTACATTCGCGTCGTC	727
OY	421	GGCGCTCCCTTAAAGGGGGCGCTGCCAGGGCCCTGGCGCATGCGCTCGGGATTCGTGAGGAC	480
Db	728	GGCGCTCCCTTAAAGGGGGCGCTGCCAGGGCCCTGGCGCATGCGCTCGGGATTCGTGAGGAC	787
OY	481	GGCGGAACTATGCGAACAGAGGAATTAACCGGTTGCTCTTTCCTATCTTCTCTTGGCT	540
Db	788	GGCGGAACTATGCGAACAGAGGAATTAACCGGTTGCTCTTTCCTATCTTCTCTTGGCT	847
OY	541	TTGCTGTCCTGTTTGACCAATTCAGACTTCGCGTTATGAAAGTCGCAACGTGTCGGGATC	600
Db	848	TTGCTGTCCTGTTTGACCAATTCAGACTTCGCGTTATGAAAGTCGCAACGTGTCGGGATC	907
OY	601	TACCATGTCAAGACAGATTGCTCCACATCAGCATGTGTACAGAGACAGCGGACATGATC	660
Db	908	TACCATGTCAAGACAGATTGCTCCACATCAGCATGTGTACAGAGACAGCGGACATGATC	967
OY	661	ATGACACACCCCGGGGTGTGTGCGCTGTGTCGGGAGGGTAATTCCTCCGCGTGGGGATA	720
Db	968	ATGACACACCCCGGGGTGTGTGCGCTGTGTCGGGAGGGTAATTCCTCCGCGTGGGGATA	1027
OY	721	GGCGTCACTCCACGCTGCGGCGCAAGACGCGACATCCCACTGCGACAATAAGACGC	780
Db	1028	GGCGTCACTCCACGCTGCGGCGCAAGACGCGACATCCCACTGCGACAATAAGACGC	1087
OY	781	CACGTCGATTTGCTGTGTTGGGGCGGCTGCTTCTGCTCCGCTATGTACGTGGGGGATCTC	840
Db	1088	CACGTCGATTTGCTGTGTTGGGGCGGCTGCTTCTGCTCCGCTATGTACGTGGGGGATCTC	1147

QY	841	TCGGGATCTGTTTTCTCTGTCCTCAGCTGTTCACCTTCGGCCCTGGCCGACATCAAGC	900
Db	1148	TGCGGATCTGTTTTCTCTGTCCTCAGCTGTTCACCTTCGGCCCTGGCCGACATCAAGC	1207
QY	901	GTAACGAGCTCAATGTGTTCAATCTATCCCGGACAGATCAGGTCACCGGATGGCTGG	960
Db	1208	GTAACGAGCTCAATGTGTTCAATCTATCCCGGACAGATCAGGTCACCGGATGGCTGG	1266
QY	961	GATATGATGATGAATGGTCACTTACAGCAGCCCTAGTGTATGCGACGTACTCCGATC	1020
Db	1268	GATATGATGATGAATGGTCACTTACAGCAGCCCTAGTGTATGCGACGTACTCCGATC	1322
QY	1021	CCACAGCTGTGTGGGACATGATGGCGGGGCCCATGCGGGAGTCTGGCGGGCTTGGC	1080
Db	1328	CCACAGCTGTGTGGGACATGATGGCGGGGCCCATGCGGGAGTCTGGCGGGCTTGGC	1388
QY	1081	TACTATTCCATGATGGGGAATGGGCTTAAGGCTTGTGTTGATGCTACTCTTGGCCGGC	1140
Db	1388	TACTATTCCATGATGGGGAATGGGCTTAAGGCTTGTGTTGATGCTACTCTTGGCCGGC	1447
QY	1141	GTTGACGGGGAACTTTCACGACACAGGGGGGACACAGGCGCGCGCCCAAGGGCTTACA	1200
Db	1448	GTTGACGGGGAACTTTCACGACACAGGGGGGACACAGGCGCGCGCCCAAGGGCTTACA	1507
QY	1201	TCCGCTTTCACACCTGGGGCGGCTCAGAAAATCCAGCTTGTAAAACCAACGGCGAGCTGG	1266
Db	1508	TCCGCTTTCACACCTGGGGCGGCTCAGAAAATCCAGCTTGTAAAACCAACGGCGAGCTGG	1566
QY	1261	CACATCAACAGAACTGCTTGAACTGCAAATGCTCCCTCAAACTGGGTTCTTGGCCGGC	1320
Db	1568	CACATCAACAGAACTGCTTGAACTGCAAATGCTCCCTCAAACTGGGTTCTTGGCCGGC	1627
QY	1321	CTGTTTCTACAGCACAGGTTCAATGCGTCCGGATGCTCAAGCGGATGGCCAGCTGCCGC	1380
Db	1628	CTGTTTCTACAGCACAGGTTCAATGCGTCCGGATGCTCAAGCGGATGGCCAGCTGCCGC	1687
QY	1381	CCCATTTGACAGTTTCGATCAGGGGTGGGGTCCCATCACTTATTAATGATGCCACGGCTTG	1440
Db	1688	CCCATTTGACAGTTTCGATCAGGGGTGGGGTCCCATCACTTATTAATGATGCCACGGCTTG	1744
QY	1441	GACCAGAGCCCTTATGCTGGCACTACGCACTTCAACCCGTGTGGTATCGTCCCGCTTG	1500
Db	1748	GACCAGAGCCCTTATGCTGGCACTACGCACTTCAACCCGTGTGGTATCGTCCCGCTTG	1807
QY	1501	CAGGTGTGTGGCCCAAGTACTGTGTTCACTCCAAAGCCCTGTTGTGTGGGGACACACCGAT	1566
Db	1808	CAGGTGTGTGGCCCAAGTACTGTGTTCACTCCAAAGCCCTGTTGTGTGGGGACACACCGAT	1867
QY	1561	CGTTTCGGCGCCCTTACGTACAGATGGGGGTGAGATGAGACGACAGTCTCTTCTCAAC	1620
Db	1868	CGTTTCGGCGCCCTTACGTACAGATGGGGGTGAGATGAGACGACAGTCTCTTCTCAAC	1922
QY	1621	AACACGCGGCGCGCAACGGGGCAACTGGTTCGGCTGTACATGATGAAATGACACCGGGTTC	1680
Db	1928	AACACGCGGCGCGCAACGGGGCAACTGGTTCGGCTGTACATGATGAAATGACACCGGGTTC	1987
QY	1681	ACCAAGAGCTGTGGGGGCCCCCGGTGCAACATCGGGGGGGTGGCAACAACATTGATC	1740
Db	1988	ACCAAGAGCTGTGGGGGCCCCCGGTGCAACATCGGGGGGGTGGCAACAACATTGATC	2047
QY	1741	TGCCCAACGAGCTCTTCCTCGAGAACATCCGAGGCACTTTCACCAAAATGCGGTTCCGGG	1800
Db	2048	TGCCCAACGAGCTCTTCCTCGAGAACATCCGAGGCACTTTCACCAAAATGCGGTTCCGGG	2107
QY	1801	CTTGTGTTG 1809	
Db	2108	CTTGTGTTG 2116	

RESULT 2
US-09-539-601-1
; Sequence 1, Application US/09539601C
; Patent No. 6630343

```
GENERAL INFORMATION:
APPLICANT: Bartenschlager, Ralf FW
TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
FILE REFERENCE: all sequences
CURRENT APPLICATION NUMBER: US/09/539,601C
EARLIER FILING DATE: 2001-08-30
EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
NUMBER OF SEQ ID NOS: 51
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1
LENGTH: 11076
TYPE: DNA
ORGANISM: Hepatitis C virus
FEATURE:
NAME/KEY: 5'UTR
LOCATION: (1)..(341)
OTHER INFORMATION: construct 1389/Core-3'/wt
FEATURE:
NAME/KEY: CD8
LOCATION: (342)..(1193)
OTHER INFORMATION: hepatitis C virus core - neomycin
OTHER INFORMATION: phosphotransferase fusion protein
FEATURE:
NAME/KEY: RBS
LOCATION: (1302)..(1812)
OTHER INFORMATION: internal ribosome entry site from
OTHER INFORMATION: encephalomyocarditis virus
FEATURE:
NAME/KEY: CD8
LOCATION: (1813)..(10845)
OTHER INFORMATION: hepatitis C virus polypeptide from core to
OTHER INFORMATION: nonstructural protein NS5B; parental sequence
OTHER INFORMATION: without cell culture-adaptive mutations
FEATURE:
NAME/KEY: 3'UTR
LOCATION: (10846)..(11076)
US-09-539-601-1

Query Match      86.8%; Score 1570.6; DB 4; Length 11076;
Best Local Similarity 91.8%; Pred. No. 0;
Matches 1660; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

QY      1  ATGAGCAGCAATCTTAACCTCAAGAAACCAACGTAACCAACGCGCGCCACAG 60
DB      1813  ATGAGCAGCAATCTTAACCTCAAGAAACCAACGTAACCAACGCGCGCCACAG 1872
QY      61  GACGTCAAGTCCCGGCGGCGGTGATCGTGTGGAGTTTAACTGTTGCGGCGAG 120
DB      1873  GACGTCAAGTCCCGGCGGCGGTGATCGTGTGGAGTTTAACTGTTGCGGCGAG 1932
QY      121  GGGCCAGGTTGGGTGCGCGGCACTAGAGAACTTCCGAGCGGTGCAACTCGTGA 180
DB      1933  GGGCCAGGTTGGGTGCGCGGCACTAGAGAACTTCCGAGCGGTGCAACTCGTGA 1992
QY      181  AGCGCAGCACTATCCCAAGCTGCGCAGCGCCGAGGGCGAGGCGCTGCGTCA 240
DB      1993  AGCGCAGCACTATCCCAAGCTGCGCAGCGCCGAGGGCGAGGCGCTGCGTCA 2052
QY      241  TACCCCTGGCCCTCTATGAGCAAGAGGAGTGGGAGAGGAGTGGGCTCGTCA 300
DB      2053  TACCCCTGGCCCTCTATGAGCAAGAGGAGTGGGAGAGGAGTGGGCTCGTCA 2112
QY      301  CGTGGCTCCCGGCTAGTGGGCGCCCACTGACCCCGGCGTAGTCTGCGTAAT 360
DB      2113  CGTGGCTCCCGGCTAGTGGGCGCCCACTGACCCCGGCGTAGTCTGCGTAAT 2172
QY      361  AAGTATGATGATCCTCTACATGCGGCTTGGCGCACTTATGGGGTATCAT 420
DB      2173  AAGTATGATGATCCTCTACATGCGGCTTGGCGCACTTATGGGGTATCAT 2232
QY      421  GGGCTCCCTTATGAGGCGGCGTCCAGAGGCGCTGCGCATGCGGCTTGA 480
DB      421  GGGCTCCCTTATGAGGCGGCGTCCAGAGGCGCTGCGCATGCGGCTTGA 480
```

```
DB      2233  GGGCGCCCCCTTAGGGGGCGCTGCGAGGCGCTGCGCATGAGGCTCGGAGTTGAGAGAC 2292
QY      481  GGGCTGAATATGCAACAGGAAATTTACCGGTTGCTTTCTCTATCTTCTTGGGT 540
DB      2293  GGGCGAATATGCAACAGGAAATCTGCGGTTGCTCTTTTCTATCTTCTTGGGT 2352
QY      541  TTGCTGCTGTTTGGCAATTCAGCTTCGCTTATGAGTGGCAAGCTGTCGGGATC 600
DB      2353  TTGCTGCTGTTTGGCAATTCAGCTTCGCTTATGAGTGGCAAGCTTATCCGAGTG 2412
QY      601  TACCATGTCAGCAAGATTTCTCACTCAAGCATGTGTACAGACAGCGGACATGATC 660
DB      2413  TACCATGTCAGCAAGATTTCTCACTCAAGCATGTGTATGAGGACAGGACATGATC 2472
QY      661  ATGACACACCCCGGGTGTGTGCTGTGCGGGAGGTAATTTCTTCCGCTGCTGGT 720
DB      2473  ATGACATACCCCGGGTGTGTGCTGTGCGGGAGAACATCTCTCCGCTGCTGGGTA 2532
QY      721  GGGCTCACTCCGACGCTGCGGCGCAAGAGCGCAGATCCCACTGCGCAATACGACG 780
DB      2533  GGGCTCACTCCGACGCTGCGGCGCAAGAGCGCTAGGCTCCCACTACAGATACGACG 2592
QY      781  CAGTCGATTTGCTGTGGGCGGCTGCTTCTGCTCGCTATGATCCTGCGGAGATCTC 840
DB      2593  CAGTCGATTTGCTGTGGGCGGCTGCTTCTGCTCGCTATGATCCTGCGGAGATCTC 2652
QY      841  TGGGATCTGTTTCTGCTGCTCTCACTGCTTCACTTCTGCTGCGGCAATCAGAGC 900
DB      2653  TGGGATCTGTTTCTGCTGCTCTCACTGCTTCACTTCTGCTGCGGCAATCAGAGC 2712
QY      901  GTACAGGATGCAATTTGTTCAATCTATCCCGGCGAGTACAGTCAACGAGGCTTGG 960
DB      2713  GTACAGGATGCAATTTGTTCAATCTATCCCGGCGAGTACAGTCAACGAGTACGCTTGG 2772
QY      961  GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
DB      2773  GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2832
QY      1021  CCAAGAGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
DB      2833  CCAAGAGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2892
QY      1081  TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
DB      2893  TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2952
QY      1141  GTTGAAGGGAAGCTTACAGACAGAGGGAACAAGCGCGCGCGCCACGCGGCTTACA 1200
DB      2953  GTTGAAGGGAAGCTTACAGACAGAGGGAACAAGCGCGCGCGCCACGCGGCTTACA 3012
QY      1201  TCCCTCTTCAACTGCGCGGCTCAAGAAATTCAGCTTGTAAACCAACGAGCTGG 1260
DB      3013  TCCCTCTTCAACTGCGCGGCTCAAGAAATTCAGCTTGTAAACCAACGAGCTGG 3072
QY      1261  CACATCAAGAGAGCTGCTTGAATGCAATGATCCTCCCAACTGGGTTCTTGGCGG 1320
DB      3073  CACATCAAGAGAGCTGCTTGAATGCAATGATCCTCCCAACTGGGTTCTTGGCGG 3132
QY      1321  CTGTTCTACACGCAAGTTCATGCGTCCGATGCTCAGAGCGATGCGCAGCTGCGC 1380
DB      3133  CTGTTCTACACGCAAGTTCATGCGTCCGATGCTCAGAGCGATGCGCAGCTGCGC 3192
QY      1381  CCCATTGACCAAGTTCATGCGTCCGATGCTCAGAGCGATGCGCAGCTGCGC 1440
DB      3193  CCCATTGACCAAGTTCATGCGTCCGATGCTCAGAGCGATGCGCAGCTGCGC 3252
QY      1441  GACAGAGGCGCTATGCTGAGCACTACGCACTCAACGCTGAGTATGCTGCGGCTGG 1500
DB      3253  GACAGAGGCGCTATGCTGAGCACTACGCACTCAACGCTGAGTATGCTGCGGCTGG 3312
QY      1501  CAGGTGTGAGCCAGTGTACTGTTCACTCAAGCCCTGTGTGTGTGGAGCAACGAT 1560
DB      3313  CAGGTGTGAGCCAGTGTACTGTTCACTCAAGCCCTGTGTGTGTGGAGCAACGAT 3372
```


QY 781 CACGTGATTTGCTGTTGGGGGCGCTGCTTCTGCTCCGCTATGATGACGTGGGGGATCTC 840
DB 2593 CATGTGATTTGCTGTTGGGGGCGCTGCTTCTGCTCCGCTATGATGACGTGGGGGATCTC 2652
QY 841 TCGCGATCTGTTTCTGCTGCTCTCTCACTGCTGTTTCACTTCTCGCTCGCGCATCATAGACG 900
DB 2653 TCGCGATCTGTTTCTGCTGCTCCGCTGTTTCACTTCTCGCTCGCGCATCATAGACG 2712
QY 901 GTACAGGATCTGCAATGTTCAATCTATCCCGGCACATGATCAGGTACACGGCATGCTTGG 960
DB 2713 GTACAGGATCTGCAATGTTCAATCTATCCCGGCACATGATCAGGTACACGGCATGCTTGG 2772
QY 961 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
DB 2773 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2832
QY 1021 CCACAAAGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
DB 2833 CCACAAAGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2892
QY 1081 TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
DB 2893 TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2952
QY 1141 GTTGAACGGGGAACTTTACACGACAGGGGGGACACACGCGCGCGCCACGCGGCTTACA 1200
DB 2953 GTTGAACGGGGAACTTTACATGATGACAGGGGGGACAGATGGCCAAAAACCCCTCGGGA 3012
QY 1201 TCCCTCTTCAACCTGCGCGCGCGCTCAAGAAATTCAGCTGTTAAACACCAACGCGAGCTG 1260
DB 3013 TCCCTCTTCAACCTGCGCGCGCTCAACAAATTCAGCTGTTAAACACCAACGCGAGCTG 3072
QY 1261 CACATCAACAGAACTGCTTGAATGCAATGCACTCCCTCAAACTGGGTTCTTGGCCGCG 1320
DB 3073 CACATCAACAGAACTGCTTGAATGCAATGCACTCCCTCAAACTGGGTTCTTGGCG 3132
QY 1321 CTGTTCTACACGACAGGATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
DB 3133 CTGTTCTACACGACAGGATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3192
QY 1381 CCCATTCACAGGATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
DB 3193 CCCATTCACAGGATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3252
QY 1441 GACCAAGAGGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
DB 3253 GACCAAGAGGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3312
QY 1501 CAGGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
DB 3313 CAGGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3372
QY 1561 CGTTTCGCGCGCTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
DB 3373 CGTTTCGCGCGCTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3432
QY 1621 AACACGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
DB 3433 AACACGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3492
QY 1681 AACACGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
DB 3493 AACACGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3552
QY 1741 TGCCCCAAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
DB 3553 TGCCCCAAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3612
QY 1801 CTTTGGTTG 1809
DB 3613 CTTTGGTTG 3621

RESULT 5
US-09-539-601-31
Sequence 31, Application US/09539601C
Patent No. 6630343
GENERAL INFORMATION:
APPLICANT: Barteneschlager, Ralf FW
TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
FILE REFERENCE: all sequences
CURRENT APPLICATION NUMBER: US/09/539,601C
CURRENT FILING DATE: 2001-08-30
EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
EARLIER FILING DATE: 1999-04-03
NUMBER OF SEQ ID NOS: 51
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 31
LENGTH: 11076
TYPE: DNA
ORGANISM: Hepatitis C virus
FEATURE:
NAME/KEY: 5'UTR
LOCATION: (1)..(341)
OTHER INFORMATION: construct 1389/Core-3/19
FEATURE:
NAME/KEY: CDS
LOCATION: (342)..(1193)
OTHER INFORMATION: hepatitis C virus core - neomycin
OTHER INFORMATION: phosphotransferase fusion protein
FEATURE:
NAME/KEY: RBS
LOCATION: (1202)..(1812)
OTHER INFORMATION: internal ribosome entry site from
OTHER INFORMATION: encephalomyocarditis virus
FEATURE:
NAME/KEY: CDS
LOCATION: (1813)..(10845)
OTHER INFORMATION: hepatitis C virus polyprotein from core to
OTHER INFORMATION: nonstructural protein NS5B; carries cell culture
OTHER INFORMATION: adaptive mutations from clone no. 19
FEATURE:
NAME/KEY: 3'UTR
LOCATION: (10846)..(11076)
US-09-539-601-31
Query Match 86.8%; Score 1570.6; DB 4; Length 11076;
Best Local Similarity 91.8%; Pred. No. 0; Mismatches 149; Indels 0; Gaps 0;
Matches 1660; Conservative 0; Mismatches 149; Indels 0; Gaps 0;
QY 1 ATGAGCAGCAATCTCTAAACCTCAAAAGAAACCAAACTGAACCAACCGCGCCACAG 60
DB 1813 ATGAGCAGCAATCTCTAAACCTCAAAAGAAACCAAACTGAACCAACCGCGCCACAG 1872
QY 61 GACGTCAAGTTCCCGGCGGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 120
DB 1873 GACGTCAAGTTCCCGGCGGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1932
QY 121 GGGCCCAAGTTGGGTGTGCGGCGACCTAGAGAACTTCCGAGCGGTGCAACCTCTGGA 180
DB 1933 GGGCCCAAGTTGGGTGTGCGGCGACCTAGAGAACTTCCGAGCGGTGCAACCTCTGGA 1992
QY 181 AGGCGCAACCTATCTCCCAAGCTGCGCAGCCCGAGGCGAGGGCTGAGGCTCAAGCCGCG 240
DB 1993 AGGCGCAACCTATCTCCCAAGCTGCGCAGCCCGAGGCGAGGGCTGAGGCTCAAGCCGCG 2052
QY 241 TACCTTTGGCCCTCTATGCGCAAGAGGCGATGGGGTGTGCGAGATGCTCTGTCACC 300
DB 2053 TACCTTTGGCCCTCTATGCGCAAGAGGCGCTTGGGGTGTGCGAGATGCTCTGTCACC 2112
QY 301 CGTGGCTCCGCGCTAGTGGGGGCGCCACTGACCCCGCGGTAGGTGCGCAATTTGGGT 360
DB 2113 CGTGGCTCCGCGCTAGTGGGGGCGCCACTGACCCCGCGGTAGGTGCGCAATTTGGGT 2172
QY 361 AAAGTCATGATACCTTCATATGCGGCTGCGAGACTCATGAGGTACATTCGCTGCTC 420

Db	2173	AAGGTCATCGATACCCTCAGCGTGCGGCTTGC	CCGATCTCATGCGGTACATTC	CCGCTGCTC	2232
Qy	421	GGCGCTCCCTTAAGGGGGCGCTG	CCAGGGCCCTGGGCGCATGCGCTG	CCGGATTCTG	AGAGC 480
Db	2233	GGCGCCCCCTCAGGGGGCGCTG	CCAGGGCCCTGGGCGCATGCGCTG	CCGGATTCTG	AGAGC 2292
Qy	481	GGCGTGAACATATGAAACAGGGGAATTTAC	CCGGGTGCTCTTCTCATCTTC	CTCTGGCT	540
Db	2293	GGCGTGAACATATGAAACAGGGGAATTTAC	CCGGGTGCTCTTCTCATCTTC	CTCTGGCT	2352
Qy	541	TTGCTGTCCTGTTTGACCATTCACAGCTTC	CCGCTTATGAAGTGGCAACG	TGTCGGGATC	600
Db	2353	TTGCTGTCCTGTTTGACCATTCACAGCTTC	CCGCTTATGAAGTGGCAACG	TGTCGGGATC	2412
Qy	601	TACCATGTACGAAACGATTGCTCCA	CTCAAGCATGTTACGAGACAG	CGGACATGATC	660
Db	2413	TACCATGTACGAAACGATTGCTCCA	CTCAAGCATGTTGTTAGAGCAG	CGGACATGATC	2472
Qy	661	ATGCAACACCCCGGGGTGTGTGTCCTGTGT	CCGGGAGGGTAATTTCTCCCGCTG	CTGGGTA	720
Db	2473	ATGCAATACCCCGGGGTGTGTGTCCTGTGT	CCGGGAGGAACAATCTCCCGCTG	CTGGGTA	2532
Qy	721	GGCGTCACTCCACGCTGCGGGCCACAGGAC	GCACAGCATCCGCACTAATACAGACG		780
Db	2533	GGCGTCACTCCACGCTGCGGGCCACAGGAC	GCCTAAGCTTACGCTCCCACTAAGCATAC	AGACG	2592
Qy	781	CACGTGATTTGCTGTTGGGGCGGCTG	CCCTTGTCTCCGCTATGTACGTGGGGATCTC		840
Db	2593	CATGTGATTTGCTGTTGGGGCGGCTG	CTGTGCTCCGCTATGTACGTGGAGATCTC		2652
Qy	841	TGCGGATCTGTTTTCCTGCTCTCTC	AGCTTTCACCTTCTGCGCTCGCGCAATC	AGACG	900
Db	2653	TGCGGATCTGTTTTCCTGCTCTCTC	AGCTTTCACCTTCTGCGCTCGCGCAAGAC		2712
Qy	901	GTACAGGACGTGCATTTGTTCAATCTATC	CCGGGACACGATCAGGTCACCGATGGCTTG		960
Db	2713	GTACAGGACGTGCATTTGTTCAATCTATC	CCGGGACACGATCAGGTCACCGATGGCTTG		2772
Qy	961	GATATGATGATGAACCTGGTCACTTA	CAGCAACCCCTAGTGATGACAGCTACTC	CCGATC	1020
Db	2773	GATATGATGATGAACCTGGTCACTTA	CAGCAACCCCTAGTGATGACAGCTACTC	CCGATC	2832
Qy	1021	CCACAAGCTGCTGTGACATGCTGGCGGGG	CCCACTGGGAGTCTGCGGGGCTTGGCC		1080
Db	2833	CCACAAGCTGCTGTGACATGCTGGCGGGG	CCCACTGGGAGTCTGAGCGGGGCTTGGCC		2892
Qy	1081	TACATATTCATGATGGGGGAATCGGGCT	TAAGTCTTGGTTGTGATGTACTCTTTG	CCGGC	1140
Db	2893	TACATATTCATGATGGGGGAATCGGGCT	TAAGTCTTGGTTGTGATGTACTCTTTG	CCGGC	2952
Qy	1141	GTTCACGGGGGAACCTTACACGACAGGGG	GAACAACGCGCCGCGCCACAGGGCTTACA		1200
Db	2953	GTTCACGGGGGAACCTTATGTACAGGGGG	GAATGACCAAAACACCTCTGGGATTTACG		3012
Qy	1201	TCCCTCTTTCACACCTGGGCGGGCT	CAGAAAAATCCAGCTTGTAAACCAAC	CGGACGTGG	1260
Db	3013	TCCCTCTTTCACACCGGGGTATCC	CAGAAAAATCCAGCTTGTAAACCAAC	CGGACGTGG	3072
Qy	1261	CACATTCACAGAACTGCTTGAATCTG	CAATGATCTCCCTCAAACTGGGATCTCTTG	CCGGC	1320
Db	3073	CACATTCACAGAACTGCTTGAATCTG	CAATGATCTCCCTCAAACTGGGATCTCTTG	CCGGC	3122
Qy	1321	CTGTTTCAACAGCAAGGTTCAATGCGT	CCGGATGCTCAGAGCGCATGGCCAGCTCGCGC		1380
Db	3133	CTGTTTCAACAGCAAGGTTCAATGCGT	CCGGATGCTCAGAGCGCATGGCCAGCTCGCGC		3192
Qy	1381	CCCATTCACAGTTGATCAGGGGTGGGGT	CCCATCACTTAAATGAATGCCAGCGCTTG		1440
Db	3193	CCCATTCACAGTTGATCAGGGGTGGGGT	CCCATCACTTAAATGAATGCCAGCGCTTG		3252
Qy	1441	GACCAAGGCGCTATGCTGGCACTAC	GCACTCAACCTGTGGTATGTGCGCGCGCTTG		1500

```

Db      3253 GACCAGAGGCGCTTATTGTTGGCACTAGCAACCCCGCGTGCAGTATCGTACCCGCGGCG 3312
Qy      1501 CAGGTGTGTGCGCCAGGTACTGTGTTTACATCCAAAGCCCTGTGTGTGTGGGAGCGACGAT 1560
Db      3313 CAGGTGTGTGTCTCAGTGTACTGTCTTACACCCCAAGCCCTGTCTGTGTGTGGGAGCGACGAC 3372
Qy      1561 CGTTTCGCGCGCCCTTACGTACAGATGTGGGATGAGAATGAGACGAGCGTGTCTTCTCAAC 1620
Db      3373 CGGTTCCGCGCTCTTACGTACAGTGTGGGGAGAAATAGACGAGCGTGTCTTCTTAAAC 3432
Qy      1621 AACACGCGCGCCCGCACGAGGCGCAACTGTGTTCGGCTGTATCATGATGAATAGACCGGAGTTC 1680
Db      3433 AACACGCGCGCCCGCACGAGGCGCAACTGTGTTGGCTGTATCATGATGAATAGACACTGGAGTTC 3492
Qy      1681 ACCAAGACGTGTGTGGGGGCGCCCGCGGTGCAACATCGGGGGGGGTCCGCGCAACAACCTTTGATC 1740
Db      3493 ACCAAGACGTGTGTGGGGGCGCCCGCGGTGCAACATCGGGGGGGATCCGCAATTAACCTTTGATC 3552
Qy      1741 TGCCCCACGGAAGTGTCTTCCGGAAGCATCCGAGGCGCACTTACCAATGATGCGGTTCCGAGG 1800
Db      3553 TGCCCCACGGAAGTGTCTTCCGGAAGCATCCGAGGCGCACTTACCAAGTGTGTGTTCCGAGG 3612
Qy      1801 CCTTGCTTG 1809
Db      3613 CCTTGCTTG 3621

RESULT 6
US-09-014-416-4
; Sequence 4, Application US/09014416
; Patent No. 6153421
; GENERAL INFORMATION:
; APPLICANT: Yanagi, Masayuki
; APPLICANT: Bukh, Jens
; APPLICANT: Emerson, Susanne U.
; APPLICANT: Purcell, Robert H.
; TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 20264276
; CURRENT APPLICATION NUMBER: US/09/014,416
; CURRENT FILING DATE: 1998-01-27
; EARLIER APPLICATION NUMBER: US 60/053,062
; EARLIER FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 9595
; TYPE: DNA
; ORGANISM: Hepatitis C virus
US-09-014-416-4

Query Match      86.0%; Score 1556.2; DB 3; Length 9595;
Best Local Similarity 91.3%; Pred. No. 0;
Matches 1651; Conservative 0; Mismatches 158; Indels 0; Gaps 0;

Qy      1 ATGAGCAGCATCTCTTAACCTCAAGAAAAACCAACGTAACACCAACCGCGGCCACAG 60
Db      342 ATGAGCAGCATCTCTTAACCTCAAGAAAAACCAACGTAACACCAACCGCGGCCACAG 401
Qy      61 GACGTCAAGTTCCTCGGCGCGGTGTGTCAGATGTTGTGTGAAGTTTACCTGTTCGCGGCAAG 120
Db      402 GACGTCAAGTTCCTCGGCGCGGTGTGTCAGATGTTGTGTGAAGTTTACCTGTTCGCGGCAAG 461
Qy      121 GGCCCGAAGTTGGGTGTGCGCGGCACTTAGAGAACTTCGAGCGGTTGCGCAACTCGTGA 180
Db      462 GGCCCGAAGTTGGGTGTGCGCGGCACTTAGAGAACTTCGAGCGGTTGCGCAACTCGTGA 521
Qy      181 AGCGGACAACCTATCCCAAGGCTCGCCAGCCCGAGGAGCAAGGCGCTGAGCCCGG 240
Db      522 AGCGGACAACCTATCCCAAGGCTCGCCAGCCCGAGGAGCAAGGCGCTGAGCCCGG 581
Qy      241 TACCTTGGCCCTCTATTGGCAAGAGGCGATGGGGTGGGCAAGATGGCTCTGTACCC 300

```


Db 582 TACCTTGGCCCCCTATGAGCAATGAGGCTTGAGGAGGAGAGATGAGCTCTGTCAACC 641
Qy 301 CGTGGCTCCGGGCTAGTTGGGGCCCACTGACCCCGGCGATGAGGTATGATTTGGAGT 360
Db 642 CCGGCTCCGGGCTAGTTGGGGCCCACTGACCCCGGCGATGAGGTATGATTTGGAGT 701
Qy 361 AAGATCATGATACCTTATGATGAGGCTTTCGCGAAGCTTATGAGGAGTATTCGCTGTC 420
Db 702 AAGATCATGATACCTTATGATGAGGCTTTCGCGAAGCTTATGAGGAGTATTCGCTGTC 761
Qy 421 GCGGCTCCCTTATGAGGAGGCTTTCGCGAAGCTTATGAGGAGTATTCGCTGTC 480
Db 762 GCGGCTCCCTTATGAGGAGGCTTTCGCGAAGCTTATGAGGAGTATTCGCTGTC 821
Qy 481 GCGGCTCCCTTATGAGGAGGCTTTCGCGAAGCTTATGAGGAGTATTCGCTGTC 540
Db 822 GCGGCTCCCTTATGAGGAGGCTTTCGCGAAGCTTATGAGGAGTATTCGCTGTC 881
Qy 541 TTGCTGCTCTGTTTGAACATTCAGCTTCCGCTTATGAGTGCAGACGCTGTCGAGATC 600
Db 882 CTGCTGCTCTGTTTGAACATTCAGCTTCCGCTTATGAGTGCAGACGCTGTCGAGATC 941
Qy 601 TACCATGTCAGAACGATGCTTCACTCAGACATGCTGTCAGAACGAGACATGATC 660
Db 942 TACCATGTCAGAACGATGCTTCACTCAGACATGCTGTCAGAACGAGACATGATC 1001
Qy 661 ATGCACACCCCGGAGTGTGCTGTCGAGGAGGAGTATTCCTCCCGCTGCTGAGT 720
Db 1002 ATGCACACCCCGGAGTGTGCTGTCGAGGAGGAGTATTCCTCCCGCTGCTGAGT 1061
Qy 721 GCGCTACTCCACGCTCGCGGCAAGAGCAGACATCCCATGCGACATACAGACG 780
Db 1062 GCGCTACTCCACGCTCGCGGCAAGAGTGCAGAGTCCCATGCGACATACAGACG 1121
Qy 781 CAGCTGATTTGCTGTTGGGCGGCTGCTTCTGCTCCGCTATGATGCTGGGAGATCTC 840
Db 1122 CAGCTGATTTGCTGTTGGGCGGCTGCTTCTGCTCCGCTATGATGCTGGGAGATCTC 1181
Qy 841 TGGGATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 1182 TGGGATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1241
Qy 901 GTAACGAGCTGCAATTTGTTCAATCTATCCCGGCAAGTATCAGATCAGCGATGCTTGG 960
Db 1242 GTGCGAGATGCAATGCTGCTCAATCTATCCCGGCAAGTATCAGATCAGCGATGCTTGG 1301
Qy 961 GATATGATGATGAATGCTGCTCAATCTATCCCGGCAAGTATCAGATCAGCGATGCTTGG 1020
Db 1302 GATATGATGATGAATGCTGCTCAATCTATCCCGGCAAGTATCAGATCAGCGATGCTTGG 1361
Qy 1021 CCACAAAGCTGCTGAGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1080
Db 1362 CCACAAAGCTGCTGAGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1421
Qy 1081 TACTATTCATGAGTGGGAGACTGAGGCTTATGAGTCTTGTGATGCTACTTTTGGCGAG 1140
Db 1422 TACTATTCATGAGTGGGAGACTGAGGCTTATGAGTCTTGTGATGCTACTTTTGGCGAG 1481
Qy 1141 GTTGAAGGAGAACTTACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200
Db 1482 GTTGAAGGAGAACTTACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1541
Qy 1201 TCCCTCTTCAACCTTGGGCGGCTGAGAAATTCAGCTTGTAAACCAACGAGAGCTTGG 1260
Db 1542 TCCCTCTTCTCACTGGGCGGCTGAGAAATTCAGCTTGTGTAAATCAACGAGAGCTTGG 1601
Qy 1261 CACATCAAGAGACTGCTTGAACATGAGTCTCCCTCAAACTGGGTTCTTGGCGAG 1320
Db 1602 CACATCAAGAGACTGCTCCCTTAATTCAGATGATCTCCCTCAAACTGGGTTCTTGGCGAG 1661
Qy 1321 CTGTTTACAGCAGAGGTTCAATGCTGCTGAGTGTCTAGAGGAGGAGGAGGAGGAGGAGG 1380
Db 1662 CTGTTTACAGCAGAGGTTCAATGCTGCTGAGTGTCTGAGGAGGAGGAGGAGGAGGAGG 1721

Qy 1381 CCCATTGACAGATTGATTCAGAGGAGTGGGATCCCATCTATATATGAGTCCACGAGCTTG 1440
Db 1722 CCATTTAGACTGATTCGCGCAGAGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1781
Qy 1441 GACCAAGAGGCTTATGCTGAGCTTACAGACCTTCAACCGTGTGATTCGCTGAGCTTGG 1500
Db 1782 GATCAGAGGCTTATGCTGAGCTTACAGACCTTCAACCGTGTGATTCGCTGAGCTTGG 1841
Qy 1501 CAGGTGTGAGGCTTATGCTGAGCTTACAGACCTTCAACCGTGTGATTCGCTGAGCTTGG 1560
Db 1842 CAGGTGTGAGGCTTATGCTGAGCTTACAGACCTTCAACCGTGTGATTCGCTGAGCTTGG 1901
Qy 1561 CGTTTGGGCGGCTTACAGTACAGATGAGGAGTGAAGATGAGACGAGCTGCTTCTGAC 1620
Db 1902 CGTTTGGGCGGCTTACAGTACAGATGAGGAGTGAAGATGAGACGAGCTGCTTCTGAC 1961
Qy 1621 AACACGAGGCTTACAGTACAGATGAGGAGTGAAGATGAGACGAGCTGCTTCTGAC 1680
Db 1962 AACACGAGGCTTACAGTACAGATGAGGAGTGAAGATGAGACGAGCTGCTTCTGAC 2021
Qy 1681 AACACGAGGCTTACAGTACAGATGAGGAGTGAAGATGAGACGAGCTGCTTCTGAC 1740
Db 2022 ACTAAGAGGCTTACAGTACAGATGAGGAGTGAAGATGAGACGAGCTGCTTCTGAC 2081
Qy 1741 TGCCCCACGAGCTTCCGAGAGCATCCGAGGAGCACTTACACCAATGCGTTTCGAG 1800
Db 2082 TGCCCCACGAGCTTCCGAGAGCATCCGAGGAGCACTTACACCAATGCGTTTCGAG 2141
Qy 1801 CTTTGGTTG 1809
Db 2142 CTTTGGTTG 2150

RESULT 7
US-09-014-416-6
Sequence 6, Application US/09014416
Patent No. 6153421
GENERAL INFORMATION:
APPLICANT: Yanagi, Masayuki
APPLICANT: Buhn, Jens
APPLICANT: Emerson, Susanne U.
APPLICANT: Purcell, Robert H.
TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND
FILE REFERENCE: 20264276
CURRENT APPLICATION NUMBER: US/09/014,416
CURRENT FILING DATE: 1998-01-27
EARLIER APPLICATION NUMBER: US 60/053,062
EARLIER FILING DATE: 1997-07-18
NUMBER OF SEQ ID NOS: 65
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 6
LENGTH: 9599
TYPE: DNA
ORGANISM: Hepatitis C virus
US-09-014-416-6

Query Match 86.0%; Score 1556.2; DB 3; Length 9599;
Best Local Similarity 91.3%; Pred. No. 0;
Matches 1651; Conservative 0; Mismatches 158; Indels 0; Gaps 0;

Qy 1 ATGAGCAGAGTCCCTAAACCTCAAGAAACCAAGGTAACACCAACCGCGGCGACAG 60
Db 342 ATGAGCAGAGTCCCTAAACCTCAAGAAACCAAGGTAACACCAACCGCGGCGACAG 401
Qy 61 GACGTCAGATTCGCGGCGGCTGATCAGATCGTTGGTGAAGTTTACCTGTTGCGGAGG 120
Db 402 GACGTCAGATTCGCGGCGGCTGATCAGATCGTTGGTGAAGTTTACCTGTTGCGGAGG 461
Qy 121 GCGCCAGGTTGGGTGTCGCGGCACTAGAAAGCTTCCAGCGGTCGCAACTCTGTGGA 180
Db 462 GCGCCAGGTTGGGTGTCGCGGCACTAGAAAGCTTCCAGCGGTCGCAACTCTGTGGA 521


```

QY 181 AGGCGAACCTATCCCAAGGCTGGCGAGCCCGGAGGCGCTGGGCTGAGCCCGG 240
DB 522 AGGCGAACCTATCCCAAGGCTGGCGAGCCCGGAGGCGCTGGGCTGAGCCCGG 581
QY 241 TACCCCTGGCCCTCTATGAGCAAGAGGCGATGGGGTGGGAGAGATGCTCTGTACCC 300
DB 582 TACCCCTGGCCCTCTATGAGCAAGAGGCGCTGGGGTGGGAGAGATGCTCTGTACCC 641
QY 301 CGTGGCTCCCGGCTATGTTGGGGGCCCACTGACCCCCCGGCGTAGTGGGTAAATTTGGGT 360
DB 642 CGCGGCTCCCGGCTATGTTGGGGGCCCACTGACCCCCCGGCGTAGTGGGTAAATTTGGGT 701
QY 361 AAGATCATGATACCTCTCAATGCGGCTTCCGAGCTCAATGGGGTCAATTCGCGCTGTC 420
DB 702 AAGATCATGATACCTCTCAATGCGGCTTCCGAGCTCAATGGGGTCAATTCGCGCTGTC 761
QY 421 GCGGCTCCCTTAGGGGGCGCTGCGAGGCGCTGGCGCATGAGCGTCCGGGTTCTGAGAGAC 480
DB 762 GCGGCTCCCTTAGGGGGCGCTGCGAGGCGCTGGCGCATGAGCGTCCGGGTTCTGAGAGAC 821
QY 481 GCGGCTCACTATGCAACAGGGAATTTAACCGGTTGCTCTTTCTATCTTTCTCTGGCT 540
DB 822 GCGGCTCACTATGCAACAGGGAATTTAACCGGTTGCTCTTTCTATCTTTCTCTGGCT 881
QY 541 TTGCTGCTCTGTTTGAACAATCCAGCTTCCGCTTATGAAGTGGGCAAGCTGTCGGGATC 600
DB 882 CTGCTGCTCTGTTTGAACAATCCAGCTTCCGCTTATGAAGTGGGCAAGCTGTCGGGATC 941
QY 601 TACCATGTCAGCAACGATTTGCTCAACTCAAGCATGCTGTACAGACAGCGGACATGATC 660
DB 942 TACCATGTCAGCAACGATTTGCTCAACTCAAGCATGCTGTGTATGAGGCAAGCGGACATGATC 1001
QY 661 ATGCAACACCCCGGGGTGTGTGCTGTGTCGGGAGGATATCTTCCGCTGCTGGGTA 720
DB 1002 ATGCAACACCCCGGGGTGTGTGCTGTGTCGGGAGGATATCTTCCGCTGCTGGGTA 1061
QY 721 GGGCTCACTCCCAAGCGTCCGGGCGAAGGAGCGAGATCCCACTCGGCAATACGACGC 780
DB 1062 GGGCTCACTCCCAAGCGTCCGGGCGAAGGAGCGAGATCCCACTCGGCAATACGACGC 1121
QY 781 CACGTGATTTGCTGTTGGGGCGGCTGCTTCTGCTCGCTATGATGAGGGGATCTC 840
DB 1122 CACGTGATTTGCTGTTGGGGCGGCTGCTTCTGCTCGCTATGATGAGGGGATCTC 1181
QY 841 TGGGATCTGTTTTCTGCTCTCTCTGAGCTGTTCACTTCTCGCTCGCGGACATGACGC 900
DB 1182 TGGGATCTGTTTTCTGCTCTCTCTGAGCTGTTCACTTCTCGCTCGCGGACATGACGC 1241
QY 901 GTCAGGAGCTGCAATTTGTTCAATCTATCCCGGCAAGTACAGGTACCGCATGGCTGG 960
DB 1242 GTCAGGAGCTGCAATTTGTTCAATCTATCCCGGCAAGTACAGGTACCGCATGGCTGG 1301
QY 961 GATATGATGATGAATGTCATCTTACCTTACAGGAGCCCTAGTGGTATCGAGCTACCGGATC 1020
DB 1302 GATATGATGATGAATGTCATCTTACCTTACAGGAGCCCTAGTGGTATCGAGTTCCTCGGATC 1361
QY 1021 CCACAAGCTGTGTCGATGATGTCGCGGGGCCCACTGGGAGATCTTGGCGGCTTGGC 1080
DB 1362 CCACAAGCTGTGTCGATGATGTCGCGGGGCCCACTGGGAGATCTTGGCGGCTTGGC 1421
QY 1081 TACTATTCATGATGGGGAATCTGGGCTTAAAGTCTTGGTGTATGCTACTTTTGGCGGC 1140
DB 1422 TACTATTCATGATGGGGAATCTGGGCTTAAAGTCTTGGTGTATGCTACTTTTGGCGGC 1481
QY 1141 GTTGAACGGGGAACCTTACAGCAGAGGGGGAACACAGCGCGCCGCCCAAGGGCTTACA 1200
DB 1482 GTTGAACGGGGAACCTTACAGCAGAGGGGGAAGGTGGCGGCAACACCTCCGGGTTACG 1541
QY 1201 TCCCTTTTCAACACTTGGGCGGCTCAAAAATCAGCTTGTATAACCAACGCGACGCTGG 1260
DB 1542 TCCCTTTTCAACACTTGGGCGGCTCAAAAATCAGCTTGTATAACCAACGCGACGCTGG 1601

```

```

QY 1261 CACATCAACAGAACTGCTTGAATGCAATGATCCTCTCAAACTGGGTTCTTCCGCGG 1320
DB 1602 CACATCAACAGAACTGCTTGAATGCAATGATCCTCTCAAACTGGGTTCTTCCGCGG 1661
QY 1321 CTGTTCTTACAGCAGAGTTCAATGCTTCGAGTGTCAAGCGCATGCGCAGCTGCGC 1380
DB 1662 CTGTTCTTACAGCAGAGTTCAATGCTTCGAGTGTTCAGCGCTTCAGCGTGTCTTACCGGTTG 1841
QY 1381 CCCATTGACCAATTTGATAGAGGGGTGGGTCCCATCACTTAATGATGCCACGCGTTG 1440
DB 1722 CCCATTGATGATTTGGCCAGGGGGTGGGGCCCATCACTTAATGATGCCATTAACAGCTGG 1781
QY 1441 GACCAAGGCGCTTATTTGCTGGCACTACGACCTCAACCGTGTGATCGTGCCTGGGTTG 1500
DB 1782 GATCAGAGGCTTATTTGCTGGCACTACGACCTTCAGCGTGTGTGTCTTACCGGTTG 1841
QY 1501 CAGGTGTGTGCGCCAGTGTACTGTTTCACTCCAGCGCTGTTGTGGTGGGACGACGAT 1560
DB 1842 CAGGTGTGTGCGCCAGTGTACTGTTTCACTCCAGCGCTGTTGTGGTGGGACGACGAT 1901
QY 1561 CGTTTGGCGCCCTTACGTAAGATGGGTGAGATGAGACGGAAGTGTGCTTCTCAAC 1620
DB 1902 CGTTTGGCGCCCTTACGTAAGATGGGTGAGATGAGACGGAAGTGTGCTTCTCAAC 1961
QY 1621 AACAGCGGCGCGCAAGGGGCAACTGTTGCGCTGTACATGATGATGATGACAGCGGTTTC 1680
DB 1962 AACAGCGGCGCGCGCAAGGGGCAACTGTTGCGCTGTACATGATGATGATGATGACAGCGGTTTC 2021
QY 1681 ACCAAGACGTTGGGGGCGCCCGGTGCAACATCGGGGGGTCGCAACCAACTTTGATC 1740
DB 2022 ACTAAGACGTTGGGGGCGCCCGGTGTAACATCGGGGGGTCGTAACCGCACTTTGATC 2081
QY 1741 TGGCCCAAGGATGCTTTCGGAAGCATCCCGAGGCGCACTTAACCAAAATGCGGTTGCGGG 1800
DB 2082 TGGCCCAAGGATGCTTTCGGAAGCATCCCGAGGCGCACTTAACCAAAATGCGGTTGCGGG 2141
QY 1801 CCTGGTTG 1809
DB 2142 CCTGGTTG 2150

RESULT 8
US-09-827-688-6
; Sequence 6, Application US/09827688
; Patent No. 6821955
; GENERAL INFORMATION:
; APPLICANT: ORSON, FRANK
; APPLICANT: KINSEY, BERNIA
; APPLICANT: BHOGLI, BALBIR
; TITLE OF INVENTION: MACROAGGREGATED PROTEIN CONJUGATES AS ORAL GENETIC IMMUNIZATION
; FILE REFERENCE: P01949US1/10004014
; CURRENT FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/195,680
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 9413
; TYPE: DNA
; ORGANISM: HEPATITIS C
US-09-827-688-6

Query Match 85.3%; Score 1543.4; DB 4; Length 9413;
Best Local Similarity 90.8%; Pred. No. 0;
Matches 1643; Conservative 0; Mismatches 166; Indels 0; Gaps 0;

QY 1 ATGAGCAGCAATCTTAACTCAAGAAACCAAAAGTAAACCAACGCGCGCCCAAG 60
DB 330 ATGAGCAGCAATCTTAACTCAAGAAACCAAAAGTAAACCAACGCGCGCCCAAG 389
QY 61 GACGTCAAGTCCCGGCGGTGTAGATGTTGTGAGATTAACTGTTGCCGCGCAG 120

```

Db 390 GACGTTAAGTCCCGGCGGTGTCAGATCGTTGGAGATTACCTGTTGCGCGCAGG 449
Qy 121 GGGCCAGGTTGGGTGTGGCGCGCATAGAGACTTCGAGCGGTCCGCAACTGTGGA 180
Db 450 GGGCCAGGTTGGGTGTGGCGCGCATAGAGACTTCGAGCGGTCCGCAACTGTGGA 509
Qy 181 AGGCGCAACACTATCCCGAAGGCTCGCGCAGCGGCGAGGAGGCTGAGCCCGG 240
Db 510 AGGCGCAACACTATCCCGAAGGCTCGCGCAGCGGAGGAGGAGGCTGAGCCCGG 569
Qy 241 TACCTTGGGCTCTCTATGCGAACAAGGATGAGGATGAGGATGAGGATGAGGAT 300
Db 570 TACCTTGGGCTCTCTATGCGAACAAGGATGAGGATGAGGATGAGGATGAGGAT 629
Qy 301 GGTGCTCCCGGCTAGTTGGGGGCGCGCATGACCCCGCGGATGAGGATGAGGAT 360
Db 630 GGTGCTCCCGGCTAGTTGGGGGCGCGCATGACCCCGCGGATGAGGATGAGGAT 689
Qy 361 AAAGTCATGATACCTTCATGCGGCTTTCGCGCATGATGAGGATGAGGATGAGGAT 420
Db 690 AAAGTCATGATACCTTCATGCGGCTTTCGCGCATGATGAGGATGAGGATGAGGAT 749
Qy 421 GGGCTTCCCTTGGGGGCGGTGCGCAGGCTTGGCGCATGAGGCTTGGAGAGAC 480
Db 750 GGGCTTCCCTTGGGGGCGGTGCGCAGGCTTGGCGCATGAGGCTTGGAGAGAC 809
Qy 481 GGGCGAATATGCAACAGGGGAATTACCGGGTGTCTTCTCTATGCTTCTGCTGCT 540
Db 810 GGGCGAATATGCAACAGGGGAATTACCGGGTGTCTTCTCTATGCTTCTGCTGCT 869
Qy 541 TTGCTGTCTTGTGTTACCATTCACGCTTCCGTTATGAGTGGGACGAGTCCGGATC 600
Db 870 TTGCTGTCTTGTGTTACCATTCACGCTTCCGTTATGAGTGGGACGAGTCCGGATC 929
Qy 601 TACCATGTCAGAGATGTTGCTTCACTTACGATGCTGTGTCAGAGACAGGAGATGATC 660
Db 930 TACCATGTCAGAGATGTTGCTTCACTTACGATGCTGTGTCAGAGACAGGAGATGATC 989
Qy 661 ATGCAACCGCGGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 990 ATGCAACCGCGGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1049
Qy 721 GGGCTCACTCCCGCGCGCGCGCAAGAGCGCGAGATCCCGATGCGACATACGAGCG 780
Db 1050 GGGCTCACTCCCGCGCGCGCGCAAGAGCGCGAGATCCCGATGCGACATACGAGCG 1109
Qy 781 CACGTGATTTGCTGTTGGGGGCGGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 1110 CACGTGATTTGCTGTTGGGGGCGGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1169
Qy 841 TGGCGATGTTTCTGCTGCTCTCTAGAGCTTCACTTCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 1170 TGGCGATGTTTCTGCTGCTCTCTAGAGCTTCACTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1229
Qy 901 GTACAGAGATGCAATGTTCAATCTATCCCGCGCATGATGAGTCACTGAGGATGAGT 960
Db 1230 GTACAGAGATGCAATGTTCAATCTATCCCGCGCATGATGAGTCACTGAGGATGAGT 1289
Qy 961 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
Db 1290 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1349
Qy 1021 CCACAGAGCT 1080
Db 1350 CCACAGAGCT 1409
Qy 1081 TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Db 1410 TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1469
Qy 1141 GTTACGCGGGAACCTTACAGACAGGCGGAGACACAGGCGCGCGCCACGCGGCTTACCA 1200

Db 1470 GTTACGCGGCAACCCACATGACAGGCGGAGAGGATGAGCTTCCAGACCCAGAGCTCTG 1529
Qy 1201 TCCCTTTACACCTGGGCGGGCTCAGAAATCCAGCTTGTAAACACCAAGGAGCTG 1260
Db 1350 TCTGGCTTCAAGAGGCCATCTCAGAAATCCAGCTTGTAAACACCAAGGAGCTG 1589
Qy 1261 CACATCAACAGAGCTGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
Db 1590 CACATCAACAGAGCTGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1649
Qy 1321 CTGTTTCAACGACAGGTTCAATGCTTCCGATGCTTCAAGGCGATGAGGCGATGAGGCTG 1380
Db 1650 CTGTTTCAACGACAGGTTCAATGCTTCCGATGCTTCAAGGCGATGAGGCGATGAGGCTG 1709
Qy 1381 CCATGACAGGTTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
Db 1710 CCATGACAGGTTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1769
Qy 1441 GACAGAGGCGCTTATGCTGACATGACGACCTCAACCGTGTGATGATGATGATGATGATGATGATGATGATGAT 1500
Db 1770 GACAGAGGCGCTTATGCTGACATGACGACCTCAACCGTGTGATGATGATGATGATGATGATGATGATGATGAT 1829
Qy 1501 CAGGTGTGAGCCAGTGTATGTTTCACTCAAGCCCTGTTGTGTGAGGAGCAAGCAT 1560
Db 1830 CAGGTGTGAGCCAGTGTATGTTTCACTCAAGCCCTGTTGTGTGAGGAGCAAGCAT 1889
Qy 1561 GGTTCGCGCGCCCTACGATCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
Db 1890 GGTTCGCGCGCCCTACGATCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1949
Qy 1621 AACAGCGGCGCGCAACGCGGCACTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
Db 1950 AACAGCGGCGCGCGCAACGCGGCACTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2009
Qy 1681 ACAGAGCTGTGGGGGCGCGCGCGCAATCGCGGGGAGTGGGCAACAACATTTGATC 1740
Db 2010 ACAGAGCTGTGGGGGCGCGCGCGCAATCGCGGGGAGTGGGCAACAACATTTGATC 2069
Qy 1741 TGCGCCAGGAGCTGTTCCGAGAGATCCGAGGCACTTACACCAATGCGGTTGCGGG 1800
Db 2070 TGCGCCAGGAGCTGTTCCGAGAGATCCGAGGCACTTACACCAATGCGGTTGCGGG 2129
Qy 1801 CTTGCTTG 1809
Db 2130 CCGTGGTG 2138

RESULT 9
US-08-612-973-49
; Sequence 49, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEBERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHAYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973

```

FILING DATE: 11-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 2433 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2430
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..2427
US-08-612-973-49:

Query Match      85.1%; Score 1538.6; DB 3; Length 2433;
Best Local Similarity 90.7%; Pred. No. 0;
Matches 1640; Conservative 0; Mismatches 169; Indels 0; Gaps 0;

QY 1 ATGAGCAGCAATCTTAACCTCAAGAAAACCAACGATACCAACCGCGCCCAACG 60
DB 1 ATGAGCAGCAATCTTAACCTCAAGAAAACCAACGATACCAACCGCGCCCAACG 60
QY 61 GACGTCAAGTCCCGGCGGTGTGAGATCGTGTGTGAGATTACCTGTTGCCGCGAG 120
DB 61 GACGTCAAGTCCCGGCGGTGTGAGATCGTGTGTGAGATTACCTGTTGCCGCGAG 120
QY 121 GGGCCCCAGGTGGGTGTGGCGCGAGCTAGAGAACTTCCGAGCGGTGCAACTCGTGA 180
DB 121 GGGCCCCAGGTGGGTGTGGCGCGAGCTAGAGAACTTCCGAGCGGTGCAACTCGTGA 180
QY 181 AGGCGCAACTATCCCAAGGCTGCGCAAGCCGAGGGGAGGGGCTGAGCCCGGG 240
DB 181 AGGCGCAACTATCCCAAGGCTGCGCAAGCCGAGGGGAGGGGCTGAGCCCGGG 240
QY 241 TACCTTGGCCCTCTATGAGCAAGAGGAGTGGGTGGGCAAGATGCTCTGTCAACC 300
DB 241 TACCTTGGCCCTCTATGAGCAAGAGGAGTGGGTGGGCAAGATGCTCTGTCAACC 300
QY 301 CGTGGCTCCCGGCTTATGTTGGGGGCCCACTGACCCCGCGGTATGTCGGTAATTTGGGT 360
DB 301 CGTGGCTCCCGGCTTATGTTGGGGGCCCACTGACCCCGCGGTATGTCGGTAATTTGGGT 360
QY 361 AAGATGATGATACCTCTCATGATGGGCTTGCAGCACTCATGAGGGATCAATCCGCTGCTC 420
DB 361 AAGATGATGATACCTCTCATGATGGGCTTGCAGCACTCATGAGGGATCAATCCGCTGCTC 420
QY 421 GGCAGCTCCCTTATGGGGGCGCTGCAAGGCGCTGGCGCATGAGCTCCGGGTTCTGAGAGAC 480
DB 421 GGCAGCTCCCTTATGGGGGCGCTGCAAGGCGCTGGCGCATGAGCTCCGGGTTCTGAGAGAC 480
QY 481 GGGGTGAATATGCAACAGGGAATTTAACCGGTTGCTCTTCTATCTTCTCTTGGCT 540
DB 481 GGGGTGAATATGCAACAGGGAATTTAACCGGTTGCTCTTCTATCTTCTCTTGGCT 540
QY 541 TGGGTGCTGTTGCAATTCAGCTTCGCTTATGAGTGGGCAACGTCGCGGATC 600
DB 541 TGGGTGCTGTTGCAATTCAGCTTCGCTTATGAGTGGGCAACGTCGCGGATC 600
QY 601 TACCATGTACGAAACGATGCTCCAACTCAAGCATGTGTATGAGGACAGCGACATGATC 660
DB 601 TACCATGTACGAAACGATGCTCCAACTCAAGCATGTGTATGAGGACAGCGACATGATC 660

```

```

QY 661 ATGCAACACCCCGGGGTGTGACCTGTGTGCGGGAGGGTAATTCCCGGCTGTGGGTA 720
DB 661 ATGCAACACCCCGGGGTGTGACCTGTGTGCGGGAGAACACTTTCGCGCTGGGTA 720
QY 721 GGGCTCACTCCACAGCTGCGGCGCAAGAGCGCAGATCCCACTGCGCAAAATACAGCG 780
DB 721 GGGCTCACTCCACAGCTGCGGCGCAAGAGCGCAGATCCCACTGCGCAAAATACAGCG 780
QY 781 CAGGTGATTTGCTGTGGGGCGGCTGCTTCTGTCTCGCTATGTATGATGGGGATCTTC 840
DB 781 CAGGTGATTTGCTGTGGGGCGGCTGCTTCTGTCTCGCTATGTATGATGGGGATCTTC 840
QY 841 TGGGATCTGTTTCTGCTCTGCTCATGAGCTGTACCTTCTGCGCTGCGGATACAGAG 900
DB 841 TGGGATCTGTTTCTGCTCTGCTCATGAGCTGTACCTTCTGCGCTGCGGATACAGAG 900
QY 901 GTACAGACTGCAATTTGTTCAATCTATCCCGGCGCAAGTATCAGTACCGCATGCTTGG 960
DB 901 GTACAGACTGCAATTTGTTCAATCTATCCCGGCGCAATACAGGATACCGTATGGCTTGG 960
QY 961 GATATGATGATGAATGCTGACCTTACAGAGCCTTAGTGTATTCGAGCTACTCGGATC 1020
DB 961 GATATGATGATGAATGCTGACCTTACAGAGCCTTAGTGTATTCGAGCTACTCGGATC 1020
QY 1021 CCACAAAGCTGTGTGGAATGATGGTGGCGGGGCGCACTGGGGAGTCTGCGGGGCTTGGC 1080
DB 1021 CCACAAAGCTGTGTGGAATGATGGTGGCGGGGCGCAATGGGAGTCTGCGGGGCTTGGC 1080
QY 1081 TACTATTCATGATGGGGAATGAGCTGAGTCTTGTGTGTATGATCTTCTTGGCGGC 1140
DB 1081 TACTATTCATGATGGGGAATGAGCTGAGTCTTGTGTGTATGATCTTCTTGGCGGC 1140
QY 1141 GTTGAAGGGAACCTTTACAGACAGAGGGGAGACACAGGCGCGCCGCAAGGCTTACA 1200
DB 1141 GTTGAAGGGAACCTTTACAGACAGAGGGGAGACAGGCGCGCCGCAAGGCTTACA 1200
QY 1201 TCCCTCTTACACCTGGGGCGGCTGAGAAATTCAGCTTGTAAACCAACGCGAGCTGG 1260
DB 1201 TCCCTCTTACACCTGGGGCGGCTGAGAAATTCAGCTTGTGTAAACCAACGCGAGCTGG 1260
QY 1261 CACATCAACAGAACTGCTTGAATGCAATGACTCCCTCAAACTGGGTTCTTGGCGCG 1320
DB 1261 CACATCAACAGAACTGCTTGAATGCAATGACTCCCTCAAACTGGGTTCTTGGCGCG 1320
QY 1321 CTATCTCAAAACAATAATTCAACTGCTGATGCCAGACGCTTGGCCAGCTGTCCG 1380
DB 1321 CTATCTCAAAACAATAATTCAACTGCTGATGCCAGACGCTTGGCCAGCTGTCCG 1380
QY 1381 CCCATTGACCAATTTGATAGGGGGTGGGTCCCATCACTTATATGATGCCACGGCTTG 1440
DB 1381 CCCATTGACCAATTTGATAGGGGGTGGGTCCCATCACTTATATGATGCCACGGCTTG 1440
QY 1441 GACCAAGGCGCTTATTTGCTGAGCACTACGCACTCAACCGTGTATGTCGCGCGTTG 1500
DB 1441 GACCAAGGCGCTTATTTGCTGAGCACTACGCACTCAACCGTGTATGTCGCGCGTTG 1500
QY 1501 CAGGTGTGTGCGCCAGGTACTGTTCACTCCAGCCCTGTTGTGGTGGAGAGACCAT 1560
DB 1501 CAGGTGTGTGCGCCAGGTACTGTTCACTCCAGCCCTGTTGTGGTGGAGAGACCAT 1560
QY 1561 CGTTTGGCGCCCTTACATGATGAGTGGGTGAGATGAGACGAGTGTCTTCTCAAC 1620
DB 1561 CGTTTGGCGCCCTTACATGATGAGTGGGTGAGATGAGACGAGTGTCTTCTCAAC 1620
QY 1621 AACACGCGGCGCGCAAGGGGCAACTGTTTGGCTGATCAATGATGATGACCGGGTTC 1680
DB 1621 AACACGCGGCGCGCGCAAGGGGCAACTGTTTGGCTGATCAATGATGATGACCGGGTTC 1680
QY 1681 ACCAAGACGTGTGGGGGCGCCCGGTGCAACATCGGGGGGTGCGCAACAACTTTGATC 1740
DB 1681 ACCAAGACGTGTGGGGGCGCCCGGTGCAACATCGGGGGGTGCGCAACAACTTTGATC 1740

```

QY 1741 T6CCCCAGGACTGCTTCGGAAGCATCCGAGCCACTTACACCAAAATGCGGTTGCGGG 1800
DB 1741 T6CCCCAGGACTGCTTCGGAAGCATCCGAGCCACTTACACCAAAATGCGGTTGCGGG 1800
QY 1801 CCTTGTTG 1809
DB 1801 CCTTGTTG 1809

RESULT 10

US-08-927-597-49
; Sequence 49, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BOYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2433 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2430
; NAME/KEY: mat_peptide
; LOCATION: 1..2427
US-08-927-597-49

Query Match 85.1%; Score 1538.6; DB 3; Length 2433;
Best Local Similarity 90.7%; Pred. No. 0;
Matches 1640; Conservative 0; Mismatches 169; Indels 0; Gaps 0;

QY 1 ATGACGACGAATCTCTAAACCTCAAGAAAAACCAACGTAACACCAACGCGGCCACAG 60
DB 1 ATGACGACGAATCTCTAAACCTCAAGAAAAACCAACGTAACACCAACGCGGCCACAG 60
QY 61 GACGTCAAGTTCCCGGGCGGTGTCAGATCGTTGTGAGATTAACTCTGTCGCGCAGG 120

DB 61 GACGTCAAGTTCCCGGGCGGTGTCAGATCGTTGTGAGATTAACTCTGTCGCGCAGG 120
QY 121 GGGCCCAAGTTGGGTGTGTGCGGCGACTAGAGAACTTCCGAGCGGTGCAACCTGTGGA 180
DB 121 GGGCCCAAGTTGGGTGTGTGCGGCGACTAGAGAACTTCCGAGCGGTGCAACCTGTGGA 180
QY 181 AGGCGACAACTATCCCAAGGCTGCGCAGCCGAGGGCAGAGGCTCGAGCCCGGG 240
DB 181 AGGCGACAACTATCCCAAGGCTGCGCAGCCGAGGGCAGAGGCTCGAGCCCGGG 240
QY 241 TACCTTTGGCCCTCTATAGCAAGAGGCAATGAGGAGATGAGTCTCTGTCACCC 300
DB 241 TACCTTTGGCCCTCTATAGCAAGAGGCAATGAGGAGATGAGTCTCTGTCACCC 300
QY 301 GGTGCTCCCGGCGCTAGTGGGGGCGCCACTGACCCCGGCGTAGTGGCGTAATTTGGGT 360
DB 301 GGTGCTCCCGGCGCTAGTGGGGGCGCCACTGACCCCGGCGTAGTGGCGTAATTTGGGT 360
QY 361 AAGGTCAATGATACCTTCAATGCGGCTTCCGCACTCGTGGGGTACATTCCGCTGTC 420
DB 361 AAGGTCAATGATACCTTCAATGCGGCTTCCGCACTCGTGGGGTACATTCCGCTGTC 420
QY 421 GGGCTCTCCCTTAAAGGAGGCGCTGCCAGGCGCTGCGCATGCGCTCCGAGTTCGAGAGAC 480
DB 421 GGGCTCTCCCTTAAAGGAGGCGCTGCCAGGCGCTGCGCATGCGCTCCGAGTTCGAGAGAC 480
QY 481 GGGCGTAATGATGCAAGAGGAAATTTACCGGTTGCTTTCTCTATCTTCTTGAGCT 540
DB 481 GGGCGTAATGATGCAAGAGGAAATTTACCGGTTGCTTTCTCTATCTTCTTGAGCT 540
QY 541 TTGCTGTCTGTGATGACCATTCAGCTTCCGCTTATGAGTGGCGCAAGTGTCCGGATC 600
DB 541 TTGCTGTCTGTGATGACCATTCAGCTTCCGCTTATGAGTGGCGCAAGTGTCCGGATC 600
QY 601 TACCATGTACGAACGATTTGCTCAACTCAAGCATCTGTATGAGACGCGACATGATC 660
DB 601 TACCATGTACGAACGATTTGCTCAACTCAAGCATCTGTATGAGACGCGACATGATC 660
QY 661 ATGCAACCCCCGGGGTGTGTCCTGTGCGGAGGATAATCTCCCGTGTGGGTA 720
DB 661 ATGCAACCCCCGGGGTGTGTCCTGTGCGGAGGATAATCTCCCGTGTGGGTA 720
QY 721 GCGCTCACTCCCAAGCTGCGCGGCAAGAGCAGCAATCCCACTGCGACATATGACGCG 780
DB 721 GCGCTCACTCCCAAGCTGCGCGGCAAGAGCAGCAATCCCACTGCGACATATGACGCG 780
QY 781 CACGTGATTTGCTGTGGGGCGGCTGCTTGTGCTCCGCTATGATGATGAGGATGTC 840
DB 781 CACGTGATTTGCTGTGGGGCGGCTGCTTGTGCTCCGCTATGATGATGAGGATGTC 840
QY 841 TGGCGATCTGTTTCCCTGCTCTACAGTGTACCTTCTGCGCCGCAATCAACAGC 900
DB 841 TGGCGATCTGTTTCCCTGCTCTACAGTGTACCTTCTGCGCCGCAATCAACAGC 900
QY 901 GTGAGATCTGCTTCTGCTGCTTCCAGCTGTTCACATCTGCTTCCGCGGCAATGAGAGC 960
DB 901 GTGAGATCTGCTTCTGCTGCTTCCAGCTGTTCACATCTGCTTCCGCGGCAATGAGAGC 960
QY 961 GATATGATGATGAATCTGTCCTTACGACAGCCCTTGTGATGCACTCTCGGATC 1020
DB 961 GATATGATGATGAATCTGTCCTTACGACAGCCCTTGTGATGCACTCTCGGATC 1020
QY 1021 CCAGAACTGTGTGAGACATGTGTGCGGGGCGCCACTGAGGAGTCTGCGGGCTTGCC 1080
DB 1021 CCAGAACTGTGTGAGACATGTGTGCGGGGCGCCACTGAGGAGTCTGCGGGCTTGCC 1080
QY 1081 TACTATTCAATGTGTGGGAACTGAGGCTTGTGTTGATGATCTTCTTGGCGGC 1140
DB 1081 TACTATTCAATGTGTGGGAACTGAGGCTTGTGTTGATGATCTTCTTGGCGGC 1140
QY 1141 GTTGAAGGGAACCTTACAGACAGGAGGAGACACAGCGCGCGCCACGAGGCTTACA 1200

```

Db      1141  GTGCAAGGGGCAATCCCGGTGTGCAAGAGGGGAGAGAGCTCCGATATACAGGGGCGTTGTG 1200
Qy      1201  TCCTCTTTCACACCTGTGGGCGGGCTTCAGAAAATTCAGCTTGTAAACCAACGCGACGTGG 1260
Db      1201  TCCCTCTTTTACCCCGGGGTCCGGCTTCAGAAAATTCAGACTCTGTAAACACCAACGGGCAATTGG 1260
Qy      1261  CACATTCACACAACTGGCTTGAACCTGCAGATGACATCCCTCCAAACCTGGGTTCTCTGCGGCG 1320
Db      1261  CACATTCACAGAGACTGGCTCTGAACCTGCAGACATCTCCCTCCAAACAGGGTTCTTTGCGGCA 1320
Qy      1321  CTGTTTCTACACGCAAGGTTCAATATGCGTCCGGATGCTCAGAGCGCATGGCCAGCTGGCGC 1380
Db      1321  CTATTTCTAACAAACACAAATTTCAACTCGTCTGTGATGTCACAGAGCGCTTGGCCAGCTGTGCG 1380
Qy      1381  CCAATTGACCACTTGCATCAGGGGTGGGGTCCCATCACTTAATTAATGACTCCACAGGCTTG 1440
Db      1381  TCCATTCGCAAGTTTGTGCTCAGGGGTGGGGGTCCCTCACTTACACTGAGCCTTAACAGCTCG 1440
Qy      1441  GACCAAGAGGCCCTAATTGCTGTGGCACTAGSCACTCAACCGTGTGATTCGTGCGCGCTTG 1500
Db      1441  GACCAAGAGGCCCTAATGCTGTGGCACTAGCGCCTCGAACCGTGTGATTTGTAACCGCGCTCT 1500
Qy      1501  CAGGTGTGTGGCCCAAGTGTACTGTGTTTCACTCCAAGCCCTGTTGTGTGGAGGACACGAT 1560
Db      1501  CAGGTGTGTGGCTCCAGTGTATTGTCTTCAACCCCGAGCCTGTTGTGTGGAGGACACGAT 1560
Qy      1561  CGTTTCCGGCGCCCTTACGTACGATGGGGGTGAGATGAGACGGAAGTGTCTTCTCAAC 1620
Db      1561  CGGTTTGGTGTGCCACGTATTAATCGGGGGGCGAAACGATCGGATGTGCTGATTTCTCAAC 1620
Qy      1621  AACACGCGGCGCGCGCGCGGACGACCTGGTTCGTGTATGATGAAATGAGTGCACCTGGGTTTC 1680
Db      1621  AACCAAGAGTGTGGGGGGGCCCCCGGCAACATCGGGGGGGTCCGCAACACACTTGTATC 1740
Qy      1681  ACCAAGAGTGTGGGGGGGCCCCCGGCAACATCGGGGGGGGCCGGCAACACACTTGTATC 1740
Db      1681  ACCAAGAGTGTGGGGGGGCCCCCGGCAACATCGGGGGGGGCCGGCAACACACTTGTATC 1740
Qy      1741  TGCCCAACGAGACTGTTCCGAGAGCATCCGAGGCACTTACACCAATGCGGTTCCGGGG 1800
Db      1741  TGCCCAACTGACTGTTTTCGAGAGCAGCCCGAGGCACTTACGCCAGATGGCGTTCCGGG 1800
Qy      1801  CCTGTGTTG 1809
Db      1801  CCTGTGCTG 1809

RESULT 11
US-08-150-204E-96
; Sequence 96, Application US/08150204E
; Patent No. 6538126
GENERAL INFORMATION:
APPLICANT: CHO, Joong Myung
LEE, Yong Beom
PARK, Young Woo
LIM, Kook Jin
CHOI, Deog Young
SO, Hong Seob
KIM, Chun Hyung
KIM, Sung Taek
YANG, Jae Young
TITLE OF INVENTION: HEPATITIS C DIAGNOSTICS AND VACCINES
NUMBER OF SEQUENCES: 128
CORRESPONDENCE ADDRESS:
ADDRESSEE: YANG, Jae Young
STREET: 386-1, Doryong-dong, Yuseong-gu
CITY: Daejeon
STATE: Daejeon
COUNTRY: Republic of Korea
ZIP: 305-340
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5inch, 1.44MB storage
COMPUTER: IBM PC/entium

```

```

? OPERATING SYSTEM: Windows
? SOFTWARE: Microsoft Word
?
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/150,204E
? FILING DATE: 20-Apr-1994
? CLASSIFICATION: <Unknown>
?
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: KR 91-9510
? FILING DATE: 10-JUN-1991
? APPLICATION NUMBER: KR 91-13601
? FILING DATE: 6-AUG-1991
? ATTORNEY/AGENT INFORMATION:
? NAME: Shahan Islam, Esq.
? REGISTRATION NUMBER: 32,507
? REFERENCE/DOCKET NUMBER: 2695/FLK
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212) 940-8564
? TELEFAX: (212) 940-8776
?
? INFORMATION FOR SEQ ID NO: 96
?
? SEQUENCE CHARACTERISTICS:
? LENGTH: 9472 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
?
? MOLECULE TYPE: DNA
?
? FEATURE:
?
? OTHER INFORMATION: KRCV-LBCL, Fig. 2
?
? SEQUENCE DESCRIPTION: SEQ ID NO: 96
?
? US-08-150-204E-96
?
?
? Query Match      84.6%; Score 1531.2; DB 4; Length 9472;
? Best Local Similarity 90.4%; Pred. No. 0;
? Matches 1635; Conservative 0; Mismatches 173; Indels 0; Gaps 0
?
?
? 1 ATGAGACGAAATCTTAAACCTCAAGAAAAACCAACGTAACACCAACCGCGCCACAG 60
? 343 ATAGAGACAAATCTTAAACCTCAAGAAAAACCAACGTAACACCAACCGCGCCACAG 402
? 61 GACGTCAGATTCCTCCGGGCGGTGTGCAGATCGTTGGTGAAGTTTACTCTGTTCCGCGCAG 120
? 403 GAAATTAAAGTTCCTCCGGGCGGTGTGCAGATCGTTGGTGAAGTTTACTCTGTTCCGCGCAG 462
? 121 GCGCCCAAGTTGGGTGTGTGCGCGCACTAGAAAGACTTCCGAGCGGTGCAACTCTGTGA 180
? 463 GCGCCCAAGTTGGGTGTGTGCGCGCACTAGAAAGACTTCCGAGCGGTGCAACTCTGTGA 522
? 181 AGCGCAACAACATATCCCAAGGCTGCCAGCCCAAGGCGAGGGCTGAGGCTCAAGCCCGG 240
? 523 AGCGCAACAACATATCCCAAGGCTGCCAGGCGCCCAAGGCGAGGCTGAGGCTCAAGCCCGG 582
? 241 TACCTCTGGACCTCTATAGCAACGAGGCAATGGGATGGGAGATGGCTCTGTCAACC 300
? 583 TACCTCTGGACCTCTATAGCAACGAGGCAATGGGATGGGAGATGGCTCTGTCAACC 642
? 301 CGTGAGTCCCGGCTAGTGGGAGCCCACTGACCCCGGCGTAGAGTGCAGTAATTGGGT 360
? 643 CGCGGCTCCCGGCTAGTGGGAGCCCACTGACCCCGGCGTAGAGTGCAGTAATTGGGT 702
? 361 AAGTATGATGATACCTTCAATGCGGCTTGCAGCACTCAATGAGGGATCAATTCGCGCTGC 420
? 703 AAGTATGATGATACCTTCAATGCGGCTTGCAGCACTCAATGAGGGATCAATTCGCGCTGC 762
? 421 GGGGCTCCCTTAGGGGGCGGTGCGAGGGCCCTGGCGCATGGCGTCCGGGTTCTGAGAGAC 480
? 763 GGGGCTCCCTTAGGGGGCGGTGCGAGGGCCCTGGCGCATGGCGTCCGGGTTCTGAGAGAC 822
? 481 GGGGTAACATATGCAACGAGGAATTTACCCGAGTGTCTTTCTCTATCTTCTCTTGGCT 540
? 823 GGGGTAACATATGCAACGAGGAATTTACCCGAGTGTCTTTCTCTATCTTCTCTTGGCT 882
? 541 TTGCTGTCTGTGTAACCAATTCAGCTTCGCTTATGAAAGTGGCAACGTGTCCGGGATC 600
? 883 CTGCTGTCTGTGTAACCAATTCAGCTTCGCTTATGAAAGTGGCAACGTGTCCGGGATG 942

```

QY 601 TACCAATGTCACGAGATTGCTCAACTCAAGCATGCTGTACAGACAGCGGACATGATC 660
 DB 943 TACCAATGTCACGAGATTGCTCAACTCAAGCATGCTGTGTATGAGGACGAGCATGATC 1002
 QY 661 ATGCAACACCCCGGGGTGTGTGCTGTGTGCTGGGAGGTAATTCCTCCGCTGTGGGTA 720
 DB 1003 ATGCAACACCCCGGGGTGTGTGCTGTGTGCTGGGAGGTAATTCCTCCGCTGTGGGTA 1062
 QY 721 GGGCTCACTCCACGCTGCGGGCAAGAGGCGAGAGTCCCACTGCGGCAATAGAGGCG 780
 DB 1063 GCACTTACTCCACGCTGCGGGCAAGAGGCGAGAGTCCCACTGCGGCAATAGAGGCG 1122
 QY 781 CACGTCGATTTGCTGCTGGGGCGGCTGCTTGTGCTCCGCTATGATGATGAGGATCTC 840
 DB 1123 CATTGTGACTGCTGCTGGGGTAGCTGCTTCTGTGCTGCTATGATGATGAGGAGCTCTC 1182
 QY 841 TCGGATCTGTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 DB 1183 TCGGATCTGTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1242
 QY 901 GTACAGGACTGCAATTTGTTCAATCTATCCGGGCAAGTATCAGGTCAACGCAATGAGCTTG 960
 DB 1243 GTACAGGACTGCAATTTGTTCAATCTATCCGGGCAAGTATCAGGTCAACGCAATGAGCTTG 1302
 QY 961 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
 DB 1303 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1362
 QY 1021 CCACAAAGCTGTGTGTGACATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
 DB 1363 CCACAAAGCTGTGTGTGACATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1422
 QY 1081 TACTATTCATGCTGT 1140
 DB 1423 TACTATTCATGCTGT 1482
 QY 1141 GTTGAAGGGAACCTTACACAGAGGGGAGACACAGCGCGCGCGCCACAGGAGCTTACA 1200
 DB 1483 GTTGAAGGGAACCTTACACAGAGGGGAGACACAGCGCGCGCGCCACAGGAGCTTACG 1542
 QY 1201 TCCCTCTTCAACCTGCGCGGCTGCAAAATTCAGCTTTGTAACCAACAGGAGCTTG 1260
 DB 1543 TCCCTCTTCAACCTGCGCGGCTGCAAAATTCAGCTTTGTAACCAACAGGAGCTTG 1602
 QY 1261 CACATCAACAGAGCTGCGCTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1320
 DB 1603 CACATCAACAGAGCTGCGCTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1662
 QY 1321 CTGTTCTACACGACAGGTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
 DB 1663 CTGTTCTACACGAGTTCAAGGTTCAAGGTTCAAGGTTCAAGGTTCAAGGTTCAAGGTTCA 1722
 QY 1381 CCCATTGACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
 DB 1723 CCCATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1782
 QY 1441 GACGAGAGGCGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
 DB 1783 GATCAAGAGGCGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1842
 QY 1501 CAGGTGTGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
 DB 1843 CAGGTGTGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1902
 QY 1561 CGTTTCGCGCGCCCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
 DB 1903 CGTTTCGCGCGCCCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1962
 QY 1621 AACACGCGCGCGCCCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
 DB 1963 AACGCGCGCGCGCCCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2022

QY 1681 ACCAAGAGTGTGGGGGCCCCCGGTGCAACATCGGGGGGTGCGCAACACACTTTGATC 1740
 DB 2023 ACCAAGAGATGTGGGGGCCCCCGGTGTAACATCGGGGGGTGCGCAACATACCTTGACC 2082
 QY 1741 TGGCCACGAGACTGCTTCCGAGAGCATCCGAGGCACTTACACCAATGCGGTTCCGGG 1800
 DB 2083 TGGCCACGAGACTGCTTCCGAGAGCAACCCCGGGCCACTTACACCAATGCGGTTCCGGG 2142
 QY 1801 CCTTGATT 1808
 DB 2143 CCTTGATT 2150
 RESULT 12
 US-08-324-977-11
 ; Sequence 11, Application US/08324977
 ; Patent No. 5747339
 ; GENERAL INFORMATION:
 ; APPLICANT: OKAYAMA, Hiroto
 ; APPLICANT: FUKU, Isao
 ; APPLICANT: MORI, Chisato
 ; APPLICANT: TAKAMIZAWA, Akahisa
 ; APPLICANT: YOSHIDA, Iwao
 ; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
 ; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
 ; NUMBER OF SEQUENCES: 50
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSER: Armstrong, Westernman, Hattori, Mclelland &
 ; ADDRESSER: Naughton
 ; STREET: 1725 K St. N.W. Suite 1000
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20006
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/324,977
 ; FILING DATE: 18-OCT-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 2-167466
 ; FILING DATE: 25-JUN-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 2-230921
 ; FILING DATE: 31-AUG-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 2-305605
 ; FILING DATE: 09-NOV-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/099,706
 ; FILING DATE: 30-JUL-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/769,996
 ; FILING DATE: 02-OCT-1991
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/635,451
 ; FILING DATE: 28-DEC-1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Stevens-Smith, Theresa M.
 ; REGISTRATION NUMBER: 36,281
 ; REFERENCE/DOCKET NUMBER: 900703D
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 659-2930
 ; TELEFAX: (202) 887-0357
 ; TELEX: 440142
 ; INFORMATION FOR SEQ ID NO: 11:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 6039 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: cDNA from genomic RNA
FEATURES:
NAME/KEY: misc_feature
LOCATION: 1..6039
OTHER INFORMATION: /note: "sequence = 333 - 6371 of
OTHER INFORMATION: SEQ ID NO: 1"
FEATURE:
NAME/KEY: CDS
LOCATION: 1..6039
US-08-324-977-11

Query Match 83.9%; Score 1517.8; DB 1; Length 6039;
Best Local Similarity 89.9%; Pred. No. 0;
Matches 1627; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

QY 1 ATGAGCAGCATCTTAACCTCAAGAAAACCAACGTAACCAACCGCCGCCAG 60
DB 1 ATGAGCAGCATCTTAACCTCAAGAAAACCAACGTAACCAACCGCCGCCAG 60

QY 61 GAAGTCAGATCCCGGCGGCTGTCAGATCTTGTGTGAGTTTACCTGTTCCGCGCAGG 120
DB 61 GAAGTCAGATCCCGGCGGCTGTCAGATCTTGTGTGAGTTTACCTGTTCCGCGCAGG 120

QY 121 GGGCCAGGTTGGGTGGCGCGCATGAGAACTTCGAGGGGTGCAACCTCGTGA 180
DB 121 GGGCCAGGTTGGGTGGCGCGCATGAGAACTTCGAGGGGTGCAACCTCGTGA 180

QY 181 AGCGCAGCAACTTATCCCAAGGCTCGCAGCGGAGGAGGAGGCTGAGCCGAG 240
DB 181 AGCGCAGCAACTTATCCCAAGGCTCGCAGCGGAGGAGGAGGCTGAGCCGAG 240

QY 241 TAACCTTGGCCCTCTATGAGCAAGAGGCTGAGGAGGAGGAGGCTGAGCC 300
DB 241 TAACCTTGGCCCTCTATGAGCAAGAGGCTGAGGAGGAGGAGGCTGAGCC 300

QY 301 CGGAGTCCCGGCTGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
DB 301 CGGAGTCCCGGCTGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360

QY 361 AAGATCATGATACCTCATAGCTGAGGCTGAGGAGGAGGAGGAGGAGGAGG 420
DB 361 AAGATCATGATACCTCATAGCTGAGGCTGAGGAGGAGGAGGAGGAGGAGG 420

QY 421 GGGCTCTCCCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
DB 421 GGGCTCTCCCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480

QY 481 GGGCTAACTATGCAAGGAAATTTACCGGCTTCTCTATCTTCTCTTGGCT 540
DB 481 GGGCTAACTATGCAAGGAAATTTACCGGCTTCTCTATCTTCTCTTGGCT 540

QY 541 TTGCTGTCTGTTTGAACATTCAGCTTCCGTTATGAGGAGGAGGAGGAGG 600
DB 541 TTGCTGTCTGTTTGAACATTCAGCTTCCGTTATGAGGAGGAGGAGGAGG 600

QY 601 TATCATGTGCAAGAACTGCTCAAGCAAGCATTTGTATGAGGAGGAGGAGG 660
DB 601 TATCATGTGCAAGAACTGCTCAAGCAAGCATTTGTATGAGGAGGAGGAGG 660

QY 661 ATGCAACACCCCGGAGTGTGCTGTGTGCTGAGGAGGAGGAGGAGGAGG 720
DB 661 ATGCAACACCCCGGAGTGTGCTGTGTGCTGAGGAGGAGGAGGAGGAGG 720

QY 721 GGGCTCACTCCCAAGCTCGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 780
DB 721 GGGCTCACTCCCAAGCTCGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 780

QY 781 CAGCTGATTTGCTGTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840
DB 781 CAGCTGATTTGCTGTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840

QY 841 TGGGATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900

DB 841 TGGGATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900

QY 901 GTAAGAGTGTCAATTTGTTCAATCTATCCCGGAGGAGGAGGAGGAGGAGG 960
DB 901 GTAAGAGTGTCAATTTGTTCAATCTATCCCGGAGGAGGAGGAGGAGGAGG 960

QY 961 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 961 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020

QY 1021 CCACAGCTGTGCTGAGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1080
DB 1021 CCACAGCTGTGCTGAGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1080

QY 1081 TACTATTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
DB 1081 TACTATTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140

QY 1141 GTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200
DB 1141 GTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200

QY 1201 TCCCTCTTCAACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
DB 1201 TCCCTCTTCAACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260

QY 1261 CACATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320
DB 1261 CACATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320

QY 1321 CTGTTTCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380
DB 1321 CTGTTTCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380

QY 1381 CCCATGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
DB 1381 CCCATGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440

QY 1441 GACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1500
DB 1441 GACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1500

QY 1501 CAGGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560
DB 1501 CAGGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560

QY 1561 CGTTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1620
DB 1561 CGTTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1620

QY 1621 AACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1680
DB 1621 AACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1680

QY 1681 ACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1740
DB 1681 ACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1740

QY 1741 TGGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1800
DB 1741 TGGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1800

QY 1801 CCTTGGTGG 1809
DB 1801 CCTTGGTGG 1809

RESULT 13
US-08-384-616-11
; Sequence 11, Application US/08384616
; Patent No. 5847101
; GENERAL INFORMATION:

APPLICANT: OKAYAMA, Hiroto
APPLICANT: FUKU, Isao
APPLICANT: MORI, Chisato
APPLICANT: TAKAMIZAWA, Akahisa
APPLICANT: YOSHIDA, Iwao
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Armstrong, Westerman, Hattori, Mclelland &
STREET: 1725 K St. N.W. Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/384,616
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Stevens-Smith, Theresa M.
REGISTRATION NUMBER: 36,281
REFERENCE/DOCKET NUMBER: 900703B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
TELEX: 440142
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 6039 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: cDNA from genomic RNA
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..6039
OTHER INFORMATION: /note: "sequence = 333 - 6371 of
OTHER INFORMATION: SEQ ID NO: 1"
FEATURE:
NAME/KEY: CDS
LOCATION: 1..6039
US-08-384-616-11

Query Match 83.9%; Score 1517.8; DB 2; Length 6039;
Best Local Similarity 89.9%; Pred. No. 0;
Matches 1627; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

QY 1 ATGAGCAGCAATCTCTAAACCTCAAGAAACCAACGTAACACCAACCGCCGACAG 60
DB 1 ATGAGCAGCAATCTCTAAACCTCAAGAAACCAACGTAACACCAACCGCCGACAG 60
QY 61 GACGTCAAGTCCCGGGCGGTGTCAGATCGTTGTGAGATTACCTGTTGCCGCGCAG 120

DB 61 GACGTCAAGTCCCGGGCGGTGTCAGATCGTTGTGAGATTACCTGTTGCCGCGCAG 120
QY 121 GACCCAGAGTGGGTGTGCGGCGACCTAGAGAACTTCCGACCGCTCCACCTCTGGA 180
DB 121 GACCCAGAGTGGGTGTGCGGCGACCTAGAGAACTTCCGACCGCTCCACCTCTGGA 180
QY 181 AGGCGACAACCTATCCCAAGGCTGCGACGCCAGAGGCAAGGCTGGGCTCAGCCCGG 240
DB 181 AGGCGACAACCTATCCCAAGGCTGCGACGCCAGAGGCAAGGCTGGGCTCAGCCCGG 240
QY 241 TACCTTGGCCCTCTATAGCGAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGG 300
DB 241 TACCTTGGCCCTCTATAGCGAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGG 300
QY 301 CTGCGCTCCGCGCTAGTGGGGGCGCCCACTACCCCGCGCTAGTGGCGCTTAAAGG 360
DB 301 CTGCGCTCCGCGCTAGTGGGGGCGCCCACTACCCCGCGCTAGTGGCGCTTAAAGG 360
QY 361 AAAGTCATGATACCTTCATACGCGCTTCCGCGACCTCATGGGGTACATTCCGCTGTC 420
DB 361 AAAGTCATGATACCTTCATACGCGCTTCCGCGACCTCATGGGGTACATTCCGCTGTC 420
QY 421 GCGCTCCCTTAAAGGGGCGCTGCGAGGCGCTGCGCATGGCGCTTCCGCTTCCGAGAC 480
DB 421 GCGCTCCCTTAAAGGGGCGCTGCGAGGCGCTGCGCATGGCGCTTCCGCTTCCGAGAC 480
QY 481 GCGCGTGAATATGACAAGGGAATTTACCGGTTGCTTCTCATCTTCTGCTGCTGCT 540
DB 481 GCGCGTGAATATGACAAGGGAATTTACCGGTTGCTTCTCATCTTCTGCTGCTGCT 540
QY 541 TTGCTGTCTGTTGACCATTCAGCTTCCGCTTATGAGTCCGCAAGTCCGCGATC 600
DB 541 TTGCTGTCTGTTGACCATTCAGCTTCCGCTTATGAGTCCGCAAGTCCGCGATC 600
QY 601 TACCATGTCAGAACGATGCTTCCATCACTAAGCATCTGTGACGACAGCGACATGATC 660
DB 601 TACCATGTCAGAACGATGCTTCCATCACTAAGCATCTGTGACGACAGCGACATGATC 660
QY 661 ATGCAACACCCCGGGTGTGTCCTGTCGCGGAGGTAATTCCTCCGCTGCGGTA 720
DB 661 ATGCAACACCCCGGGTGTGTCCTGTCGCGGAGGTAATTCCTCCGCTGCGGTA 720
QY 721 GCGCTCACTCCAGCTGCGGCGCAAGGACGCAAGCATCCCACTGCGCAATACGACGC 780
DB 721 GCGCTCACTCCAGCTGCGGCGCAAGGACGCAAGCATCCCACTGCGCAATACGACGC 780
QY 781 CACGTGATTTGCTGTGTTGGGCGGCTGCTTGTGCTTCCGCTAATGATGAGGAGCTC 840
DB 781 CACGTGATTTGCTGTGTTGGGCGGCTGCTTGTGCTTCCGCTAATGATGAGGAGCTC 840
QY 841 TCGCGATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
DB 841 TCGCGATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
QY 901 GTACAGAGCTCAATTTGTTCAATCTATCCGCGACGATACGATACGATGATGATGATG 960
DB 901 GTACAGAGCTCAATTTGTTCAATCTATCCGCGACGATACGATACGATGATGATGATG 960
QY 961 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
DB 961 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
QY 1021 CCACAGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1080
DB 1021 CCACAGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1080
QY 1081 TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
DB 1081 TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
QY 1141 GTTACGCGGGAACCTTACACGACAGGCGGACACACGCGCGCGCGCGCGCTTACA 1200
DB 1141 GTTACGCGGGAACCTTACACGACAGGCGGACACACGCGCGCGCGCGCGCTTACA 1200


```

Db      1141 GTTGAAGGAGATCCACGATGACAGAGGAGGCGGAGCCAAACACCAACAGGCTGTG 1200
QY      1201 TCCCTCTTACACCTGAGGCGGCTCAGAAATTCAGTTGTAAACACCAACGAGCTGG 1260
Db      1201 TCCATGTTCGCAAGTGGGCGCTCTCAGAAATCAGCTTATTAACCAATGAGAGTTGG 1260
QY      1261 CACATCAACGAACTGCTTGAATCTGAATGACCTCCCAAACTGGGTTCTTTGCCGCG 1320
Db      1261 CACATCAACGAACTGCTTGAATCTGAATGACCTCTTCCAGACTGGGTTCTTTGCCGCG 1320
QY      1321 CTGTCTTACACGACAGTTTCAATGCGTCCGAGTGTCTCAGAGCGCATGGCCAGCTGCGC 1380
Db      1321 CTGTCTTACACAGATGTTTCAATGCTGCGGAGTCCCAAGGCGCATGGCCAGCTGCGC 1380
QY      1381 CCCATTGACCAAGTTTCAATGAGGAGTGGGCTCCCATCATCTTATATAGTCTCCAGCTTG 1440
Db      1381 ACCATTGACCAAGTTTCAATGAGGAGTGGGCTCCCATCATCTTATATAGTCTCCAGATCA 1440
QY      1441 GACCAAGAGGCGCTTATGCTGAGCACTACGACCTCAACCGGTGATCGTCCGCGGTTG 1500
Db      1441 GACCAAGAGGCGCTTATGCTGAGCACTACGACCTCAACCAATGATCACTGATCGCTG 1500
QY      1501 CAGGTGTGAGGCGGCTGATCTGTTTCACTCAAGCCCTGTTGTGTGAGGAGAGCGAT 1560
Db      1501 GAGGTGTGAGGCGGCTGATCTGTTTCACTCAAGCCCTGTTGTGTGAGGAGAGCGAT 1560
QY      1561 CGTTTGGGCGGCTTACATGATGAGTGGGAGTGAATGAGACGACGCTGCTTCTCAC 1620
Db      1561 CGTTTGGGCGGCTTACATGATGAGTGGGAGTGAATGAGACGACGCTGCTTCTCAC 1620
QY      1621 AACAGCGGCGGCGGCTGATCTGTTTCACTCAAGCCCTGTTGTGTGAGGAGAGCGAT 1680
Db      1621 AACAGCGGCGGCGGCTGATCTGTTTCACTCAAGCCCTGTTGTGTGAGGAGAGCGAT 1680
QY      1681 ACCAAGACGTGTGAGGAGGCGGCTGATCTGTTTCACTCAAGCCCTGTTGTGTGAGGAG 1740
Db      1681 ACCAAGACGTGTGAGGAGGCGGCTGATCTGTTTCACTCAAGCCCTGTTGTGTGAGGAG 1740
QY      1741 TGCCCAAGGAGCTGCTTCCGGAAGCATCCGAGGCGCATTTACCAACCAATGCGGTTG 1800
Db      1741 TGCCCAAGGAGCTGCTTCCGGAAGCATCCGAGGCGCATTTACCAACCAATGCGGTTG 1800
QY      1801 CCTTGCTG 1809
Db      1801 CCTTGCTG 1809

```

RESULT 14

```

US-08-904-686A-11
/ Sequence 11, Application US/08904686A
/ Patent No. 598130
/ GENERAL INFORMATION:
/ APPLICANT: OKAYAMA, Hiroto
/ APPLICANT: FUKU, Isao
/ APPLICANT: MORI, Chisato
/ APPLICANT: TAKAMIZAWA, Akahisa
/ APPLICANT: YOSHIDA, Iwao
/ TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
/ TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
/ NUMBER OF SEQUENCES: 50
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Armelton, Westerman, Hattori, Mclelland &
/ ADDRESS: Naughton
/ STREET: 1725 K St. N.W. Suite 1000
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: U.S.A.
/ ZIP: 20006
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
/ OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
/ SOFTWARE: ASCII

```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/904,686A
FILING DATE: 01-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/324,977
FILING DATE: 18-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/099,706
FILING DATE: 30-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Mclelland, Le-Nhung
REGISTRATION NUMBER: 31,541
REFERENCE/DOCKET NUMBER: 900703G
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 6039 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA from genomic RNA
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..6039
OTHER INFORMATION: /note: "sequence = 333 - 6371 of
FEATURE:
NAME/KEY: CDS
LOCATION: 1..6039
US-08-904-686A-11

```

```

Query Match      83.9%; Score 1517.8; DB 2; Length 6039;
Best Local Similarity 89.9%; Pred. No. 0;
Matches 1627; Conservative 0; Mismatches 182; Indels 0; Gaps 0;
QY      1 ATGAGCAGCAATCTTAACTCAAGAAACCAAACTGTAACCAACGCGCCGACAG 60
Db      1 ATGAGCAGCAATCTTAACTCAAGAAACCAAACTGTAACCAACGCGCCGACAG 60
QY      61 GACGTCAAGTCCCGGCGGCGGTGATCGTGTGTGAGATTACTGTTGCCGCGCAG 120
Db      61 GACGTCAAGTCCCGGCGGCGGTGATCGTGTGTGAGATTACTGTTGCCGCGCAG 120
QY      121 GGGCCCAAGTTGGGTGGCGCGCACTAGGAAGACTTCCGACCGGTGCAACTCGTGA 180
Db      121 GGGCCCAAGTTGGGTGGCGCGCACTAGGAAGACTTCCGACCGGTGCAACTCGTGA 180
QY      181 AGGCGACAACCTATCCCAAGGCTGCGCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
Db      181 AGGCGACAACCTATCCCAAGGCTGCGCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
QY      241 TACCTTGGCCCTCTATGCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
Db      241 TACCTTGGCCCTCTATGCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
QY      301 CGTGGCTCCGCGCTAGTTGGGCGCCCACTGACCCCGCGGTAGGTGCGGTATTGGGT 360

```

```

Db      ||||||| 301 CCCGCTCCCGGCTAGTGGGGCCCAAGGACCCCGGCTAGGTGGGTATTTGGGT 360
Qy      361 AAGATCATGATACCTGACATGGGCTTCCGCACTTATAGGGATATTCGCTGTC 420
Db      361 AAGATCATGATACCTGACATGGGCTTCCGCACTTATAGGGATATTCGCTGTC 420
Qy      421 GCGGCTCCCTTATAGGGGCGCTGCAAGGCGCTGCGCATGCGTCGCGGTTCTGAGAGAC 480
Db      421 GCGGCTCCCTTATAGGGGCGCTGCAAGGCGCTGCGCATGCGTCGCGGTTCTGAGAGAC 480
Qy      481 GCGGCTGATATGCAACAGGGAATTTACCGGTTGCTTTTCTATCTTCTTGGCT 540
Db      481 GCGGCTGATATGCAACAGGGAATTTACCGGTTGCTTTTCTATCTTCTTGGCT 540
Qy      541 TTGCTGCTGTTTACCAATTCAGCTTCCGTTATGAGTGCAGACGTCGCGGATC 600
Db      541 TTGCTGCTGTTTACCAATTCAGCTTCCGTTATGAGTGCAGACGTCGCGGATC 600
Qy      601 TACATATGTCAGGAAGATGCTCCAACTCAAGCATGTTATGAGACAGCGAGATATC 660
Db      601 TATCATATGTCAGGAAGATGCTCCAACTCAAGCATGTTATGAGACAGCGAGATATC 660
Qy      661 ATGCAACACCCCGGGTGTGTCCTGTGTGTCGAGAGGTAATTCCTCCGCTGCTGGTA 720
Db      661 ATGCAATACTCTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Qy      721 GGGCTACCTCCAGCGTCCGCGCAAGAGCGCAGATCCCACTGCGCATATAGAGCC 780
Db      721 GGGCTACCTCCAGCGTCCGCGCAAGAGCGCAGATCCCACTGCGCATATAGAGCC 780
Qy      781 CAGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db      781 CAGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Qy      841 TCGGATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db      841 TCGGATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Qy      901 GTCAGGATGTCATTTGTTTCAATTCATCCCGGCAAGTACAGGATCCCGATGCTGG 960
Db      901 TTAAGGATGTCATTTGTTTCAATTCATCCCGGCAAGTACAGGATCCCGATGCTGG 960
Qy      961 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
Db      961 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
Qy      1021 CCACAAAGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1080
Db      1021 CCACAAAGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1080
Qy      1081 TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
Db      1081 TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
Qy      1141 GTTGAAGGGAAGCTTTCACAGAGGAGGAGACAGCGCGCGCGCGCGCGCGCTTACA 1200
Db      1141 GTTGAAGGGAAGCTTTCACAGAGGAGGAGACAGCGCGCGCGCGCGCGCGCTTACA 1200
Qy      1201 TCCCTCTTCAACCTGAGGCGGCTCAAGAAATTCAGCTTGTAAACCAACAGGAGCTGG 1260
Db      1201 TCCCTCTTCAACCTGAGGCGGCTCAAGAAATTCAGCTTGTAAACCAACAGGAGCTGG 1260
Qy      1261 CACATCAACAGAACTGCTTGAATGCAATGATCTCCCTCCAACTGGGTTCTTGGCGGG 1320
Db      1261 CACATCAACAGAACTGCTTGAATGCAATGATCTCTCTCAACAGGAGTTTCTTGGCGGG 1320
Qy      1321 CTGTTCTACACGACAGGTTCAATGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 1380
Db      1321 CTGTTCTACACGACAGGTTCAATGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 1380
Qy      1381 CCCATTGACAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440

```

```

Db      1381 ACCATTGAAGTTGATGACAGGAGATGGGGTCCATTACTTATGCTGATGACATCA 1440
Qy      1441 GACCAAGAGCCCTATTTGCTGACATGACGACCTTAACCGTGTGATGATGTCGCGCTTG 1500
Db      1441 GACCAAGAGCCCTATTTGCTGACATGACGACCTTAACCGTGTGATGATGTCGCGCTTG 1500
Qy      1501 CAGGTGTGAGCCGATGCTGTTTCACTTCAAGCCCTGTTGTTGTTGGAGACGACGAT 1560
Db      1501 GAGGTGTGAGCCGATGCTGTTTCACTTCAAGCCCTGTTGTTGTTGGAGACGACGAT 1560
Qy      1561 CGTTTGGCGCCCTTACGTAAGATGAGTGGGTGATGAGACGACGATGCTGCTTCAAC 1620
Db      1561 CGTTTGGGTGCTCTTACGTAAGATGAGTGGGTGATGAGACGACGATGCTGCTGCTCAAC 1620
Qy      1621 AACACGGGCGCCACAGGAGGCACTGTTGGCTGTAATGATGATGATGATGATGATGATGATG 1680
Db      1621 AACACGGGCGCCACAGGAGGCACTGTTGGCTGTAATGATGATGATGATGATGATGATGATG 1680
Qy      1681 ACCAAGAGTGTGGGGGCCCCCGTGCACATCGGGGGGTTCGCAACACACTTTGATC 1740
Db      1681 ACCAAGAGATGTGGGGGCCCCCGTGTACATCGGGGGGTTCGCAACACACTTGAAC 1740
Qy      1741 TGCCCCAGACGCTGCTTCCGAGAGCATCCGAGGCACTTACACCAAAATGTGTTCGGGG 1800
Db      1741 TGCCCCAGACGCTGCTTCCGAGAGCATCCGAGGCACTTACACCAAAATGTGTTCGGGG 1800
Qy      1801 CTTGGCTG 1809
Db      1801 CTTGGCTG 1809

```

```

RESULT 15
US-09-315-850-11
; Sequence 11, Application US/09315850
; Patent No. 6217872
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westernman, Hattori, Mclelland &
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/315,850
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/904,686
; FILING DATE: 01-AUG-1997
; APPLICATION NUMBER: US 08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605

```



```

Qy 1621 AACACGCGGCCGCCACGCGGCACTGTTCGGCTGTACATGATGATAGCACCGGGTTC 1680
    |||||
Db 1621 AACACGCGGCCGCCACGCGGCACTGTTCGGCTGTACATGATGATAGCACCGGGTTC 1680
    |||||
Qy 1681 ACCAAGACGTGTGGGGGCCCCCGTGCACATCGGGGGGTTCGCAACAACCTTGATC 1740
    |||||
Db 1681 ACCAAGACATGTGGGGGCCCCCGTGTACATCGGGGGGTTCGCAACAACCTTGACC 1740
    |||||
Qy 1741 TGCCCCACGACTGCTTCGGAAGCATCCGAGGCCACTTACACCAAAATGCGGTTGCGGG 1800
    |||||
Db 1741 TGCCCCACGACTGCTTCGGAAGCATCCGAGGCCACTTACACCAAAATGTGTTGCGGG 1800
    |||||
Qy 1801 CCTTGGTTG 1809
    |||||
Db 1801 CCTTGGCTG 1809
    |||||

```

Search completed: February 19, 2005, 11:51:43
 Job time : 316.302 secs

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	137	7.6	488	1	AV755731	AV755731 AV755731
C 2	104.4	5.8	492	1	AV7558366	AV7558366 AV7558366
C 3	58	3.2	179	7	R28798	R28798 FO-26d 22
C 4	55	3.1	844	9	CNS0052P	AL056652 Drosophil
C 5	55.4	3.1	925	9	CNS0091P	AL05013 Drosophil
C 6	51.8	2.9	879	9	CNS020QC	AL209545 Tetradom
C 7	50.2	2.8	1036	9	CNS010BS	AL098770 Drosophil
C 8	49.8	2.8	925	9	CNS0091P	AL05013 Drosophil
C 9	49.4	2.7	1101	9	CNS016UE	AL107216 Drosophil
C 10	46.4	2.6	1030	9	AG126313	AG126313 Pan trogl
C 11	46	2.5	1100	9	CNS016KD	AL106855 Drosophil
C 12	45.4	2.5	1201	9	CNS016BR	AL106545 Drosophil
C 13	45.2	2.5	839	9	CNS004NB	AL054280 Drosophil
C 14	44.2	2.4	910	9	CNS006ON	AL056529 Drosophil
C 15	44.2	2.4	1101	9	CNS007BG	AL057398 Drosophil
C 16	44	2.4	932	9	CNS0072O	AL066742 Drosophil
C 17	43.6	2.4	997	9	CNS006DN	AL065132 Drosophil
C 18	43.6	2.4	1009	9	CNS010EW	AL098882 Drosophil
C 19	43.2	2.4	1058	5	BX462094	BX462094 BX462094
C 20	42.8	2.4	1071	9	CNS000EWY	AL066395 Drosophil
C 21	42.6	2.4	645	9	CNS01213	AL101589 Drosophil
C 22	42.6	2.4	932	9	CNS0072O	AL066742 Drosophil
C 23	42.6	2.4	936	4	BG852371	BG852371 1024034A0
C 24	42.2	2.3	731	6	CB651667	CB651667 OSJNB16167

[illegible]

ALIGNMENTS

FEATURES	source
LOCUS	AV755731
DEFINITION	AV755731 BM Homo sapiens CDNA clone BMFAK03 5', mRNA sequence.
ACCESSION	AV755731
VERSION	AV755731.1
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H., Gu,Y., Li,N., Qian,B., Liu,F., Qu,D., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,D., Fu,G., Ren,S., Zhong,M., Lu,G., Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and Chen,Z.
TITLE	Homo sapiens CDNA BM clones
JOURNAL	Unpublished (2000)
COMMENT	Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919(ex.45) Fax: 86-21-50801922 Email: hanzg@cnhg.sh.cn This clone is available at CHGC in Shanghai.
FEATURES	Location/Qualifiers
source	1..488
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/cd_xref="taxon:9606"
	/clone="BMFAK03"
	/tissue_type="Bone marrow"
	/cell_type="CD34+ hematopoietic stem/progenitor cell"
	/lab_host="BM25.8"
	/clone_1lb="BM"
	/note="Vector: pTir1pEx2; Site_1: sflIA, Site_2: sflIB"

QY	856	CTGCATCTCTAAGCTTTTACACCTTCCTCCGCTCGCGACATTCAGACGGTACAGGCTCAT	915
Db	412	CTTCAAGTTTCAAGCTGATTA---TCTGGGCTTAGACCAATGATTTGTGCATGAAATGCAAC	356
QY	916	TGTTCAATCTATATCCCGGCACAGTATCAAGTCAACCCATG-GCTTGGATATGATGATGAA	974
Db	355	TGCTTCATCTATTCCTGGGCGCATCATCTGGACACCGTATGAGGATGGACATGATGATGAA	296
QY	975	CTGTGCACCTTACAGCAGCCCTAAGTGTATTCGACGCTACTCCGGATCCACAAGCTGTGCT	1033
Db	295	CTGTGCTGCACCGCTGCTATGATCATGCGTATGCGCAATGCGCGTTCCTGAGTTCATCAT	236
QY	1035	GGACATGGTGGCGGGGGGCGCACTGGGGGAGTCTCTGGCGGGCCCTGTGCTATATTCATGAT	1099
Db	235	AGATATATCATCAGCGGGGCTCATCTGGGGGCGTATGTTGGCTTAGCTTCTCTATCA	176
QY	1095	GGGGAACCTGGGCTAAGCTCTTGGTTGTATGCTACTCTTTCGCGCGCTGTGACGGGGAAC	1154
Db	175	GGGAGCGTGGCGGAAGTCCGCTGTATCTTCTGTTAACCGGTGGGGTACACGCGACGAC	116
QY	1155	TTTAACGACAGAGGGGGAC-AACGCGCGCGCCGCCACGAGGCTTACATCCCTCTTACAC	1213
Db	115	CCAGGTCAACAGTGGCACTGCTAAGCGCGCGCGCAGAGTTTCTTCAGCTTGCTACCC	56
QY	1214	CTGGGCGCGGCTCAGAAATTCAGCTTGTAAACCAACGCGACGCTGG	1260
Db	55	CTGGGCGCTAAACAGAAACATCCACTCATTTANGATNATNTTGCATATGGG	9

RESULT 2				
AV758366/c				
LOCUS	AV758366	492 bp	mRNA	linear
DEFINITION	AV758366 BM Homo sapiens cDNA clone BMFAK03 5', mRNA sequence.			

VERSION	AV758366.1	GI:10916214
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	

REFERENCE
AUTHORS
Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.
1 (bases 1 to 492)
Eumaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

TITLE	Homo sapiens cDNA BM clones
JOURNAL	Unpublished (2000)
COMMENT	Contact: Zeguang Han

Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMPAKA03"
/tissue_type="bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM2.8"
/clone_id="Bm"
/note="vector: pTriplEx2, site_1: sf1A, site_2: sf1B"

```

Query Match	5.8%	Score 104.4	DB 1	Length 492
Best Local Similarity	56.1%	Pred. No. 8.7e-17		
Matches 254; Conservative	0	Mismatches 194	Indels 5	Gaps 3

QY	812	TCAGCTCGCATAGTACGTGGGGAGATCTACGCGGATCTGTTTCTCGTCTCAGCTGT	871
Db	457	TGTATCAAGCTCATTACGTGTGGAACTCTCGCTTGGGGATACGCTTGCAGCCCACTGA	398
QY	872	TCACCTCTCTCGCCTCGCGGACATCAGACGGGTACAGACCTGCATTTGTTCAATCTATCCG	931
Db	397	TTA---TCTCTCACGACGACATCTGGTTTGTGGCAAGAAATGCACATGCTCATTTACTCTGT	341
QY	932	GCACAGTATACGGTTC-AACGATAGGCTTGGGATATAGATAGAACTGGTACACTTAACGA	990
Db	340	GCTGCATCTACCTGGAGCTACAGTATGGCAATAGGCTATAGATAGAACTGGTCCGACGCCGT	281
QY	991	GCCCTAGTGGTATGCGACAGCTACTCCGAGTCCCAAGAGCTGTCTGGAATAGTGGACGGG	1050
Db	280	TTCATGATACCTGGGGTACGCCAATGGGGTTCCTGAAGTCCATATGATATCATTAAGCTGG	221
QY	1051	GCCCACTGGGAGATCTCGCGCGGCTTGCTCTATATTCATGTGGGGAACTGGGCTTAAG	1110
Db	220	GCAACACTCGCGCGTCAATGTTTGGGCTCAGCTTACTTCCAAAGCAGGGGACGTTTGGCCAAA	161
QY	1111	GTCCTGGTGTGATGACTACTCTTTCGCGGCGTTTACAGCGGGAACTTACACACAGAGGGGG	1170
Db	160	GTAAGTGTCAATCATCTGTTCACCGCTGGGTTATGTGAGCAACACAGTATCAGTGCT	101
QY	1171	ACACAGCGCGCGCGCGGCCACGGGCT-TACATCCCTTTTCAACCTGGGCGGCTCAGAA	1228
Db	100	ATCATCGCGCGGACGACAGACAGAGTTTGGTCAGGTTGTCATATAGCTGGCGCTAAGCAGTA	41
QY	1230	AATCAGCTTTAAACACCAACGCGACGCTGCA	1262
Db	40	CATCCAGCTCATGANCCTAAATGGCAATATGGANA	8

RESULT 3	
LOCUS	R28798
DEFINITION	179 bp mRNA linear EST 03-DEC-1999
	F0-266D 22 week old human fetal liver cDNA library Homo sapiens
	cDNA clone F0-266D 5', mRNA sequence.

ACCESSION	R28798
VERSION	R28798.1
KEYWORDS	GI:6514170
SOURCE	EST.
	Homo sapiens (human)

ORGANISM	Homo sapiens
Eukaryota	Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia	Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 179)
AUTHORS	Choi,S.S., Yun,U.W., Choi,E.K., Cho,Y.G., Sung,Y.C. and Shin,H.S.
TITLE	Construction of a gene expression profile of a human fetal liver by single-pass cDNA sequencing

JOURNAL Mamm. Genome 6 (9), 653-657 (1995)
MEDLINE 96081342
PUBMED 8535075
COMMENT Contact: Hee-Sup Shin

Developmental Genetics
Pohang Institute of Science & Technology
San31, Hyojaedong Pohang, 790-784 Republic of Korea
Tel.: 562-279-2281
Fax: 562-279-2199
Email: shinh@vision.postech.ac.kr
Seq primer: T3 primer.

FEATURES

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="F0-266D"
/lab_host="XJ1-blue MRF"
/clone_lib="22 week old human fetal liver cdna library"
/notes="Vector: pRiuescriptII SK(-); Site_1: EcoRI; Site_2
XhoI: The cdna library made by oligo-dT primed and
directionally cloned between 5'Exor I-XhoI sites."

```

ORIGIN

Query Match 3.2%; Score 58; DB 7; Length 179;
 Best Local Similarity 86.5%; Pred. No. 0.00024;
 Matches 64; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1223 CTCGAGAAATTCAGCTTGTAAACACCAACGAGCTGGACATCAACAGACTGCTTGA 1282
 |||||
 DB 63 CGAGAAACATTCAGCTGTAAACGCTAATATGACAGCTGGACATCAACAGACATGCTTGA 122
 |||||

QY 1283 ACTGCATGACTCC 1296
 |||||
 DB 123 ATTCATGACTGCC 136
 |||||

RESULT 4
 CDS0052P/c 844 bp DNA linear GSS 03-JUN-1999
 LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
 DEFINITION BACR1916 of RPCI-98 library from Drosophila melanogaster (fruit
 fly); genomic survey sequence.
 AL054652
 ACCESSION AL05652.1 GI:4932342
 VERSION GSS.
 KEYWORDS Drosophila melanogaster (fruit fly)
 SOURCE Drosophila melanogaster
 ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 844)
 REFERENCE Genoscope.
 AUTHORS Direct Submission
 TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
 JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - web : www.genoscope.cns.fr)
 COMMENT Determination of this BAC-end sequence was carried out as part of a
 collaboration with the Berkeley Drosophila Genome Project (BDGP).
 The BDGP is constructing a physical map of the Drosophila
 melanogaster genome using these BACs. For further information
 please see <http://www.fruitfly.org> The BDGP Drosophila
 melanogaster BAC library was prepared by Kazutoyo Osoegawa and
 Aaron Mammose in Pieter de Jong's laboratory in the Department of
 Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
 NY. The library is named RPCI-98 and was constructed by partial
 EcoRI digestion of Drosophila DNA provided by the BDGP from the
 isogenic strain y2; cn bw sp; the same strain used for the BDGP's
 p1 and EST libraries. A more detailed description of the library
 and how to order individual BAC clones, the entire library, or
 filters for hybridization from the BACPAC Resource Center can be
 found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

1. 844
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone="BACR1916"
 /clone_11b="RPCI-98"
 /note="end : TET3"

ORIGIN

Query Match 3.1%; Score 56; DB 9; Length 844;
 Best Local Similarity 21.6%; Pred. No. 0.0012;
 Matches 76; Conservative 13; Mismatches 145; Indels 0; Gaps 0;

QY 203 CTCGCCAGCCGAGGAGGCTGGAGCTGAGCCCGGATACCTTGCCCTCTATAGCA 262
 |||||
 DB 644 CCCCTCYCCBCHYBCHTSCSYVSBSCSSSSSSSSSSBSCCBCCBGGYGCY 585
 |||||

QY 263 ACAGAGGACATGGAGGAGGAGATGCTCTCTGACACCCCTGCTCCCGGCTAGTTGG 322
 |||||
 DB 584 SBVCSBSSSTBSVSGSBSTSGTGCKSSSSSBSCSSCCCTCSSYCTCYSSCC 525
 |||||

QY 323 GCCCACTGACCCCGGCTAGGTGCGCTAATTGGTAAAGTCATGATACCTTCACAT 382
 |||||

DB 524 BYCBBGCBTBSYBSCSGBTBBSBSGGKSGKSGKGBKSGKGYCYCBTCCSBS 465
 |||||
 QY 383 GCGGCTTCCGACCTCTATGAGGATACATTCCTCTCGCGCTCCCTTAGGGGGCTG 442
 |||||

DB 464 SBCCSSSSSTBSKCCGSSSTKCGSSBGGGSSSTKSGSSSTGSSSSSTKSTS 405
 |||||

QY 443 CCAGGCGCTGGGCGCATGCGCTCCGCTTCTGAGACGCGGATGACAAAGAGA 502
 |||||

DB 404 YSTSSSSSGCGTAKNTTGTGCTTCTTSSCBSSGTSBSSTTTTGTTTTTTTT 345
 |||||

QY 503 ATTACCGGCTGCTCTTCTATGCTCTTGTGCTGCTGCTTT 554
 |||||

DB 344 TTTTGTT 293
 |||||

RESULT 5
 CDS0091P 925 bp DNA linear GSS 03-JUN-1999
 LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
 DEFINITION BACR19016 of RPCI-98 library from Drosophila melanogaster (fruit
 fly); genomic survey sequence.
 AL053013
 ACCESSION AL053013.1 GI:4934461
 VERSION GSS.
 KEYWORDS Drosophila melanogaster (fruit fly)
 SOURCE Drosophila melanogaster
 ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 925)
 REFERENCE Genoscope.
 AUTHORS Direct Submission
 TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
 JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - web : www.genoscope.cns.fr)
 COMMENT Determination of this BAC-end sequence was carried out as part of a
 collaboration with the Berkeley Drosophila Genome Project (BDGP).
 The BDGP is constructing a physical map of the Drosophila
 melanogaster genome using these BACs. For further information
 please see <http://www.fruitfly.org> The BDGP Drosophila
 melanogaster BAC library was prepared by Kazutoyo Osoegawa and
 Aaron Mammose in Pieter de Jong's laboratory in the Department of
 Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
 NY. The library is named RPCI-98 and was constructed by partial
 EcoRI digestion of Drosophila DNA provided by the BDGP from the
 isogenic strain y2; cn bw sp; the same strain used for the BDGP's
 p1 and EST libraries. A more detailed description of the library
 and how to order individual BAC clones, the entire library, or
 filters for hybridization from the BACPAC Resource Center can be
 found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

1. 925
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone="BACR19016"
 /clone_11b="RPCI-98"
 /note="end : TET3"

ORIGIN

Query Match 3.1%; Score 55.4; DB 9; Length 925;
 Best Local Similarity 14.0%; Pred. No. 0.0017;
 Matches 50; Conservative 172; Mismatches 134; Indels 1; Gaps 1;

QY 128 GATTGATGTCGCGGAGTACAGAACTTCGAGGAGTGCAGAACTCTGAGAGCGAC 187
 |||||

DB 557 GSGYGGGSSGSSBSCSSSCSSSCSCBCCCCCSCSSYCCSSSSSSSKSTSBS 616
 |||||

QY 188 AACTTATCCCAAGAGCTCCGACCCGAGGAGGAGGCTGAGTACCCGGTATACCTT 247
 |||||

DB 617 CCGCCSKSKVCGTSSSSSSSTSSSTSSSTSSSTSSSSSSSSSTSSSTSSSTSS 676
 |||||

QY 248 GGCCTCTATGCAACGAGGAGGATGGGAGATGAGTCTCTGACACCCCTGAGCT 307
 |||||

Db 677 ASGSGSWAGGAGGAGTGTSSSSSSSSSSSSSSSVSSGSKSTBSGSSSSSSSSST 736
Qy 308 CCCCCCTTAAGTTGGGGCCCCCACTGACCCCCGCGGTAGTCCGGTAAATTGGTAAGTCA 367
Db 737 SSBSCSTSS 796
Qy 368 TCGATACCTCAACATGCGGCTTCGCGACCTCATGAGGAGTATTCGCTCGCGCGCTC 427
Db 797 SSDSTSTGTCSC-CCMCTCCCTTBMBCYTSTSCGSSSSSSGKGYTKCCGCGSSSTMG 855
Qy 428 CTTAGGGGGCGCTGCGGAGGCGCTGCGCATGCGCTCCGCTTCGTGAGAGCGGCG 484
Db 856 MBGTSSACSS 912

RESULT 6
CNS0200G 879 bp DNA linear GSS 01-SEP-2000
LOCUS
DEFINITION Tetradon nigroviridis genome survey sequence T7 end of clone
15806 of library G from Tetradon nigroviridis, genomic survey
sequence.

ACCESSION AL209545.1 GI:7868364
VERSION GSS; genome survey sequence.
KEYWORDS Tetradon nigroviridis
SOURCE Tetradon nigroviridis
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthopterygii; Acanthopterygii; Percomorphi; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.

REFERENCE 1
AUTHORS Roest Croillius, H., Jailton, O., Dasilva, C., Bouneau, L., Fisher, C.,
Bernot, A., Fitzames, C., Wincker, P., Brottier, P., Quetier, F.,
Sautin, W. and Weissenbach, J.
TITLE Estimate of human gene number provided by genome-wide analysis
using Tetradon nigroviridis DNA sequence
JOURNAL Nat. Genet. 25 (2), 235-238 (2000)
MEDLINE 20296633
PUBMED 10835645

REFERENCE 2
AUTHORS Roest Croillius, H., Jailton, O., Dasilva, C., Ozouf-Costaz, C.,
Fitzames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F.,
Sautin, W., Bernot, A. and Weissenbach, J.
TITLE Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetradon nigroviridis
JOURNAL Genome Res. 10 (7), 939-949 (2000)
MEDLINE 20359837
PUBMED 10899143

REFERENCE 3 (bases 1 to 879)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetradon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetradon.
FEATURES
source
1..879
Location/Qualifiers
/organism="Tetradon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="15806"
/clone_1lb="G"
/note="Genoscope sequence ID : COAG158B03LPl-end : T7"

ORIGIN

Query Match 2.9%; Score 51.8; DB 9; Length 879;
Best Local Similarity 41.0%; Pred. No. 0.016;
Matches 128; Conservative 28; Mismatches 156; Indels 0; Gaps 0;

Qy 38 GTAACACCAACCGCCGCCACAGACGTCAAGTCCCGGCGGTGATCAATGCTGTG 97
Db 503 GYBCCCCCCCCCCCCCCCCSSBSBVGGGGAGKSGGGGGGCCCCCCCCCCCCG 562
Qy 98 GAGTTTACCTTTGCGCGGAGGGCCCCAGGTTGGTGTGCGGCACTAGAAAGATT 157
Db 563 GGGGGGGGGGGCCCCCCCCCGGGGCGSSCCCGGGGSSCGCCCCCCCCCCCC 622
Qy 158 CCGAGCGGTGCAACCTGTGTAAGGAGCAACAACCTATCCCAAGCTGAGCCGAGG 217
Db 623 CCCCCGGGGGGGGGGGGGGCCCGGCCCCCCCCCCCCCCCCCCCCCGGGGSSG 682
Qy 218 GCAGGCGCTGAGCTCAGCCCGGATACCTTGGCCCCCTTATAGCAACGAGGCAATGG 277
Db 683 GGGGGCCCGGGGCCCGGGGSSCCCCCGGSGGCCCCCGGGGGCCCCGGGGCCG 742
Qy 278 GGGCAGATGAGCTCTGTACACCCCGTGTGCTCCGAGCTAGTTGGGCCCCCACT 337
Db 743 GGGCGGSCCCCCCCCCCCCCCCCCCCCCCGSSCGSGGGGGGGGGCCCCCCCC 802
Qy 338 GCGGAGGTGCG 349
Db 803 CGGGCGGGGCGC 814

RESULT 7
CNS010BS 1036 bp DNA linear GSS 26-JUL-1999
LOCUS
DEFINITION Drosophila melanogaster genome survey sequence Sp6 end of BAC
BACN03H11 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

ACCESSION AL098770.1 GI:5610381
VERSION GSS.
KEYWORDS Drosophila melanogaster (fruit fly)
SOURCE Drosophila melanogaster
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1036)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (BDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaut at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.
FEATURES
source
1..1036
Location/Qualifiers
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN03H11"
/clone_1lb="DrosBAC"
/plasmid="pBelobAC11"
/note="end : Sp6"

ORIGIN

Query Match 2.8%; Score 50.2; DB 9; Length 1036;
Best Local Similarity 26.9%; Pred. No. 0.045;
Matches 123; Conservative 116; Mismatches 215; Indels 3; Gaps 1;

Qy 1 ATGAGCAGATTCCTTAACCTCAAGAAACCAACGACCAACCGCCCAAG 60
Db 569 ATTAATAAANANANATTAANANATTAANANATTAANANATTAANANATTAANANATTA 628
Qy 61 GACGTCAAGTCCCGGGCGGTGTCAGATCGTGTGAGTTACTGTTCGCGGCGAG 120


```

Db      629 CCAACMAACACCCSSAVSMGSAANAAAAAANSAVANMGAAGAAAAAGACTGAGAGR 688
Qy      121 GGGCCCAAGTGGGTGTGCGCGGACTGAGAGAATTCCAGCGGTCCAACTCTGTGA 180
Db      689 SSSSCCGSGSVCCYCKCGCCCGCCGAAAGAAAGACGAGSAAV--CGGSAAGCASSAGCA 745
Qy      181 AGCGCAACACTTATCCCAAGGCTCGCGACCGCGAGGCGAGGCGCTGACGCCGGG 240
Db      746 MMGASCSAASARVCGCGCGSCGSCGSCGSGMGSCGAMASSSSCSASMA 805
Qy      241 TACCTTGGCCCTTATGCAACAGAGCATGGGTGGGCGATGCTCTGTCAACC 300
Db      806 AGASAGSGGCGSCGSCGSCGSCGSCGSCGSCGSCGSCGSCGSCGSCGSCGSCG 865
Qy      301 CTGTGCTCCCGGCTGATGGGCGCCCACTGACCCCGCGGTAGTGGCTAAATTTGGT 360
Db      866 SGSGSASASGSGSTCGSSSGSGCGASMAAGSGSGSVAAAGCSAAGCRBAG 925
Qy      361 AAGTATGATGATCCTTACATGCGGCTTGGCGGACTCATGGGTATCTCCGCTGTC 420
Db      926 SSGCGACNSCASMGCKCGMAAASVSSGMAAASAAAAAARASGCGSSSGSGSGSVGS 985
Qy      421 GCGGCTCCCTTACATGCGGCTGCGACAGGCGCTGCGC 457
Db      986 BBTCSARVARAGSGCGSAGCGSSSGSSSVSGSS 1022

```

RESULT 8
CNS0091P/c
LOCUS
DEFINITION
925 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR19D16 of RCT-98 library from Drosophila melanogaster (fruit
fly); genomic survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AL051013
GI:4934461
GSS.
Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 925)

REFERENCE
AUTHORS
TITLE
JOURNAL
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mosmoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source
1. 925
Location/Qualifiers

/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR19D16"
/clone_lib="RPI-98"
/note="end : TET3"

ORIGIN

Query Match 2.8%; Score 49.8; DB 9; Length 925;
Best Local Similarity 15.4%; Pred. No. 0.056;
Matches 65; Conservative 173; Mismatches 179; Indels 5; Gaps 1;

```

Qy      112 CCGCGAGAGGAGCCCAAGTGGGTGTGCGCGGACTGAGAGAATTCCAGCGGTGCA 171
Db      924 SBCSCSCSCSSBSGSSSSMTSSNSBSCSSSSBSSTSSMSSSSBSGSSSS 865
Qy      172 CTTGTGGAAGGACAACTATATCCCAAGGCTTGGCGAGGCCGAGGCGTGGGCT 231
Db      864 SGTSAACVKKMASSGCGCGGABCMCSGSSSSCCGSAARVAVKVASGAKRGGSG 805
Qy      232 CAGCCCGGATACCTTGACCCCTTATGCAACAGAGGATGGGTGGGAGGATGGCT 291
Db      804 GASHSSSSSACSSSSSSCASGWSASSSSSASRSRGGAGGAGSSSSSSA 745
Qy      292 CTGTACCCCGGCTCCCGGCTGATGGGCGCCCACTGACCCCGCGGTAGTGGCT 351
Db      744 SAGSVVSAASSSSSSGSSSVSCSVASSMSGSSSSASASSSSSSASCSGCCCT 685
Qy      352 AATTGGGTAACTATCGATACCTTCACATGCGGCTTGGCGGACTCATGGGTGAT 411
Db      684 SWSGSCSTSAWSMASSSSSSSSSSSSASASSSSSSSSSSSSGSSGSA---- 629
Qy      412 CGGCTGTCGAGGCGCTCCCTTACGAGGCGGCGGCGGCGGATGCGCGGCTT 471
Db      628 -CGSSMSGGGSSSVASSGMSVSSGSGSGGCGGCGGSSGSSGSGSGSV 570
Qy      472 CTGAGAGAGGCGCTGATGCAACAGGAAATTTACCCGTTGCTTTCTTATCTTC 531
Db      569 CSCSSGCMCRGCSAASAAAAAASGMAASCGMCKSKSGGCTNNNNNTNTNTNTNN 510
Qy      532 CT 533
Db      509 NT 508

```

RESULT 9
CNS016U6
LOCUS
DEFINITION
1101 bp DNA linear GSS 26-JUL-1999
Drosophila melanogaster genome survey sequence T7 end of BAC
BACN16G16 of DrosBAC library from Drosophila melanogaster (fruit
fly); genomic survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AL107216
GI:5626210
GSS.
Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)

REFERENCE
AUTHORS
TITLE
JOURNAL
Genoscope.
Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (BDGP) -
<http://www.edgp.ebi.ac.uk> - This Drosophila melanogaster BAC
library (DrosBAC) was made by Alain Billand at CRPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.

FEATURES
source
1. 1101
Location/Qualifiers

/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN16G16"
/clone_lib="DrosBAC"
/plasmid="pBelobAC11"

[illegible]

RESULT 10	AG126333/c	LOCUS	AG126333	1030 bp	DNA	linear	GSS 04-NOV-2001
ACCESSION	AG126333	Pan troglodytes DNA, clone: PTB-136N19.F, genomic survey sequence.					
KEYWORDS	AG126333.1	GI:16655498					
SOURCE	GSS.						
ORGANISM	Pan troglodytes (chimpanzee)						
REFERENCE	Pan troglodytes						
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.						
TITLE	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.						
JOURNAL	BAC end sequences of library PTB						
REFERENCE	Unpublished						
AUTHORS	2 (bases 1 to 1030)						
TITLE	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.						
JOURNAL	Direct Submision						
REFERENCE	Submitted (02-AUG-2001) Asso Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Stuehli-Chou, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: chimpesgsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/), Tel:81-45-503-9111, Fax:81-45-503-9170)						
TITLE	Clones are derived from the chimpanzee BAC library pPB This BAC end was generated during the Red process and may have higher chance of Clone tracking errors.						
JOURNAL							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
REFERENCE							
AUTHORS							

FEATURES	location/Qualifiers
source	1..1030
	/organism="Pan troglodytes"
	/mol_type="genomic DNA"
	/db_xref="taxon:9598"
	/clone="PTB-136N19.F"
	/sex="male"

ORIGIN	/cell_type="lymphoblast"
	/clone_lib="PTB Chimpanzee Male BAC Library"
Query Match	2.6%; Score 46.4; DB 9; Length 1030;
Best Local Similarity	38.4%; Pred. No. 0.48;
Matches 116; Conservative	0; Mismatches 186; Indels 0; Gaps 0

QY 49 CGCGGCCCAAGAGCTCAAGTTCCCGGCGGATGTCAGATTCGTTGGTGAAGTTTAACTCG 108

Db 480 CGCCACCACTCCGCGCGCCGGGNNNNNNNNNNNGGCGCGGGGGGCGCGCGNCCGNGCCGG 421

QY 109 TTGCGCGCAGAGGCGCCCGAGGTTGGTGTGTGGCGCGACTAGGAAGACTTCCGAGCGGTG 168

Db 420 GGGCGGGGGCGGGGCTCCGCGGCGGGGGGCGCGGCGGGGCGGGCGCGCGCGCGCG 361

QY 169 CAACCTCGTGGAAAGGCGCAACTTATCCCAAGGCTCGCAGCCCGAAGGCGAGGCGCTGG 228

Db 360 GGCCCCCGGNGGNGGGCGCAACAACGCCCGCGCGCCGCGGGGGGGGGCGGNGCGGG 301

QY 229 GCTCAGGCCCGGATACCTTGTGCGCCCTCTATGTGCAAGAAGCATGGGTGTGGCAAGATGG 288

Db 300 GGGGGGNNGGGGGNGGGGGGGGCGCCCGGNGGNNGGGGGGGGGGGGGGGNGGNGGG 241

QY 289 CTCCTGTACCCCGTGGCTCCCGGCTTAGTTGGGGGCCCACTGACCCCGCGGTAGGTG 348

Db 240 NNNNNNGGNNNNNNGGGNNNGGNNCNNNNNNNNNNNNNNNNNNNNNCCCCCNNNNNNNNN 181

QY 349 CG 350

Db 180 NG 179

RESULT 11	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	TITLE	JOURNAL	COMMENT
CNS016KD	CNS016KD	Drosophila melanogaster genome survey sequence SP6 end of BAC BACN16022 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.	AL106855	AL106855.1	GI:5624152	GSS.	Drosophila melanogaster (fruit fly)	Genoscope.	Direct Submission		
							Drosophila melanogaster	Submitted (23-JUL-1999)	Genoscope - Centre National de Sequencage		
							Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)			
							1 (bases 1 to 1100)	- Web : www.genoscope.cns.fr			
								Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk - This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CBPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelovBAC11.			

FEATURES	source	location/Qualifiers
1.	1100	/organism="Proscophila melanogaster" /mol_type="genomic DNA" /db_xref="taxon:7227" /clone="BACN16D2" /clone_1b="DrosBAC" /plasmid="pbeloBAC11" /note="end : Sp6"
Query Match	2.5%	Score 46; DB 9; Length 1100

NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

```
1..1101
  /organism="Drosophila melanogaster"
  /mol_type="genomic DNA"
  /db_xref="taxon:7227"
  /clone="BACR24B13"
  /clone_lib="RPCI-98"
  /note="end : TTT3"
```

ORIGIN

```
Query Match      2.4%; Score 44.2; DB 9; Length 1101;
Best Local Similarity 23.6%; Pred. No.1.9;
Matches 77; Conservative 103; Mismatches 145; Indels 1; Gaps 1;

QY 119 GGGGGCCCCAGTTGGGTGTCGGCGCACTAGAAAGACTTCCGAGCGGTCCGAACCTGCG 178
    ||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 701 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 760

QY 179 GAAGGGGAGAACCTATCCCAAGGCTGCGCAGCCCGAGGGGAGGGGCGCTGGGCTCAGCCCG 238
    ||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 761 CGGKNCNHECCCKCCCKKKKCTCKCKBCKGCKGCKCKKKYCKKKKCTTTTCTNNYYTC 820

QY 239 GGTAACCTTGGCCCTCTATGCGACGAGGGCATGGGGTGGGCGAGATGGCTCTGTCAC 298
    ||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 821 CYTTTCTCTTCTCTTTTTKKKKKMMKKMKYKKKKKKKKKKCKGCKGCKYCKKKYKKN 880

QY 299 CCGGTGGCTCCCGGCTAGTTGGGGCCCACTGACCCCGGCGTGAAGTCCGTAATTGG 358
    ||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 881 YCYBKKBCCCCCCTCBTYKTGBNBCKKKCGKSKCKSKENBKKKKKCKKKKKCKKCB 940

QY 359 GTAAGTCATGCGATACCTCAGATG--CGGCTGGCCGACCTCATGGGGTACATTCCGCTC 417
    ||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 941 KCKYGNKCBKKKKCKKCCCGGSCCKGCKCKKCKCKGKKKKCKKCKKCKKCKKCKK 1000

QY 418 GTGGGCGCTCCCTTAGGGGGCGCTGC 443
    ||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 1001 GKCCCKNKKCKCKKCKKCKKCKKCKKCKKCKKCKKCKKCKKCKKCKKCKKCKKCKK 1026
```

Search completed: February 19, 2005, 11:38:18
Job time : 6316.46 secs

THIS PAGE BLANK (USPTO)